1955

Linear models and randomized experiments

Martin B. Wilk
Iowa State College

NOTE TO USERS

LINEAR MODELS AND RANDOMIZED EXPERIMENTS

by

Martin B. Wilk

A Dissertation Submitted to the

Graduate Faculty in Partial Fulfillment of

The Requirements for the Degree of

DOCTOR OF PHILOSOPHY

Major Subject: Statistics

Approved:

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

Head of Major Department

Signature was redacted for privacy.

Dean of Graduate College

Iowa State College

1955

TABLE OF CONTENTS

I. INTRODUCTION

A. Introductory Remarks

The primary problem underlying the investigations reported upon in this thesis concerns the derivation, interpretation and application of linear statistical models to be used in connection with the analysis of randomized experiments.

Aside from the Introduction, the thesis consists of two main Parts, II and III. The emphasis in these two parts is somewhat different, but the common element is the use of a derived linear model. The concerns and contents of Parts II and III are briefly outlined in the following sections.

Divisions B, C, D of the Introduction give some general background or introductory discussions on experimental error, randomization, models and the analysis of variance. Relationships of this thesis to other published and unpublished work are discussed in Division E of this Introduction.

1. The general problem of Part II

Part II of this thesis is concerned with some aspects of the statistical analysis of randomized comparative experiments. Generally speaking, the objective of such experiments is the evaluation of relative outputs from a set of experimental stimuli. In particular, one is concerned

with the estimation of treatment contrasts and with the estimation of measures of uncertainty (error) of such estimates; with the estimation of so-called components of variance (or measures of dispersion of populations); with obtaining objective measures of the reliability of the experiment to indicate conclusions (tests of significance). The technique of the analysis of variance plays a central role in connection with these objectives.

It has become usual in the past ten or so years to base the statistical analysis (and in particular the interpretation of the analysis of variance) of randomized experiments on assumed linear models all of whose random components are taken to be normally and independently distributed.

Given a particular linear model, with particular assumptions (and meaning) attached to its components, a meaningful interpretation of the analysis of variance may be made on the basis of the expectations* of the mean squares. However, if the technique is to be generally useful in the analysis of randomized experiments then it seems clear that the model and assumptions employed cannot be arbitrary.** Further, some indication of the inadequacy of proceeding by assumption is given by the controversy (discussed below) regarding the appropriate analysis for certain fairly simple "mixed model" situations.

Thus we take the primary problem to be that of the objective development and interpretation of linear models for randomized experiments.

---

*"Expectation" is used here with the usual meaning of average value over an indefinitely large number of repetitions.

**It may be of interest here to quote from Fisher (1935a; p. 39): "... if an experiment does allow us to calculate a valid estimate of error, its structure must completely determine the statistical procedure by which this estimate is to be calculated."

Closely associated with this are such questions as: What is the role of randomization? Under what conditions can treatment effects be estimated unbiasedly? What do the analysis of variance mean squares measure in terms of the physical situation? What assumptions are necessary to make meaningful comparisons of the mean squares? What is the relation of "fixed" and "random" factors? What constitutes replication? Does the analysis of variance always give unbiased estimates of errors of estimates and of components of variation?

A further area of the general problem revolves about questions of distributional properties of criteria used in significance tests and in interval estimation, starting from the fact that randomization is employed and using, perhaps, some reasonable simplifying assumptions.

2. The scope and structure of Part II

Part II of the present thesis gives a partial answer to some of the questions raised above. The standard designs (completely randomized, randomized block, latin square and split plot) are considered in connection with rather general experimental situations, in Divisions A, B, C and D, of Part II. For each of these, a conceptual population is defined and a linear statistical model is derived whose components reflect the parameters of the conceptual population and the properties of the (randomized) experimental design and procedure. Expectations of the analysis of variance mean squares are given, based on the derived models. The fact of randomization plays a central role.

Explicit simplifying assumptions on properties of the conceptual

population are introduced to simplify the examination of the results, and also to indicate what assumptions are necessary (or unnecessary) for an unambiguous interpretation of the mean squares. Some special cases are detailed for easy reference.

The procedure used handles both analysis of variance models and components of variance models (Eisenhart models I and II) (Eisenhart 1947), as special cases of a general model.

Some attention is given explicitly to the estimation of treatment differences and to the estimation of the errors of these estimates. We consider also the estimation of components of variation, and the selection of meaningful criteria for significance tests. Some discussion of complements such as orthogonality, biases in estimation, sensitivity of various criteria for significance tests of different null hypotheses occurs in connection with particular examples.

The order of presentation is in the directions of increasing complexity of design and experimental situation, and of diminishing assumptions. It is often true that a section is a special case of its successor.

Division E of Part II is devoted to discussions of the general methods and results, and their extension and generalization. It is to be considered an integral part of the developments given in Part II.

3. The problem and scope of Part III

The general problem of Part III concerns the use of statistical methods and concepts in the elucidation of functional relationships among physical

(experimental) variables. Some background is given in Division C of this Introduction. The rather meagre literature of relevance to this problem is reviewed in Part III.

The original contribution of Part III is an approximate test for a multiplicative-type functional structure of two factors.

B. Experimental Error and the Principle of Randomization

The term experimental error has been used in many different ways by various people. Broadly speaking it is used in connection with uncertainties or heterogeneities present in experimental investigations. To a large extent the precise intent of the phrase is a matter of individual definition.

1. Some basic issues

Any attempt to come to grips with the notion of experimental error gets one involved in some basic philosophical issues such as reproducibility of events, categorization or grouping of operations, the meaning of "true value", and so on. The nuances and detail of these matters we are neither equipped nor inclined to discuss at this time. However, in the last analysis, every procedure finds its justification only in philosophy and so we will give viewpoints on certain conceptions which play an important role in the investigations presented in the sequel. The conceptions we will discuss are "treatments", "experimental units" and "true responses".

Consider some object such as a rectangular bar. Can we speak

of the "true length" of this bar? If we attempt measurement (and our instrument is sensitive enough) then repeated measurements will not agree exactly. One might then define the "true length" as being the mean over the conceptual population of possible repetitions of the measurement; or one might regard the "true length" as having an abstract reality, independent of any measurement, and regard the observations obtained as being measurements with "errors" on this "true length". Pragmatically the two viewpoints are not distinguishable in some circumstances.

Suppose now that we take one short step from immediate physical perception and consider what would happen if we dipped the bar in acid without heating, and what would happen if we heated the bar to 200° C and after cooling dipped it in acid. We cannot do both these operations. But it seems useful, and perhaps indispensable, to conceive of properties which the bar would have as a result of either of these procedures. The notion of a "conceptual true response (or property) if such and such an operation were to be carried out" would appear to be an essential feature in experimental investigations.

Similar considerations arise with respect to the meaning of a "treatment". In attempts to understand and predict natural phenomena idealizations or abstractions such as "the application of a pressure of 50 p. s. i." are used, even though we cannot control, reproduce or even recognize with exactitude. We might think of 50 p. s. i. as representative of a certain physical procedure we are using which would have, on repetition, characteristics of variability; or we may think of an actual application

as being an attempt, frustrated by inevitable variability in technique, to obtain the abstractly existent treatment of 50 p. s. i. In this example the latter outlook appears the more productive.

When we turn, however, to the problem of comparing two varieties of corn, say for yield, (and assuming that all experimental conditions concerned are "constant" for the moment) then clearly we are concerned not with differences in individual members of the varieties but in comparisons between all members. Here our two "treatments" would involve two aggregates each of many seeds. The specific observations would be on selections from within these aggregates. It would not be natural to think of an ideal seed as representing the variety, and to think of members of the aggregate as departing more or less from the ideal. It would, however, be possible to unify this situation with that in the preceding example by such a construct. For simplicity in presentation this is the formal identification of such two situations which we shall employ.*

From the point of view of modern statistics the concept of "experimental units" would appear to have a fundamental place in the derivation of knowledge by experimentation. In certain situations, "experimental units" will have some reasonable physical identity, such as samples of a solution, plots of land, mice, etc.; but often they will be more abstract as, for example, periods of time in an individual's day, an unknown set of weather conditions, the different attitudes of

*See, however, a study based on the other view under Case 6, Division B of Part II.

an individual, etc.

It is not necessarily obvious why one should distinguish experimental units from treatments. Both can act as classification bases for conceptually possible observations. The distinction lies somewhat in our attitude or interest and somewhat in our procedure. For example, individual samples of a certain type of plastic will respond differently to a treatment of 50 p.s.i., but our interest is with some "average response". Generalizing one might wish to relate dimension to pressure, and we can think of a distinct function relating these for every block of plastic. Combining these one would have dimension as a function of two variables, namely pressure and block. We are satisfied, usually, in such a situation to use some average function of only one variable as representative for all blocks. A construct would be to regard individual blocks as deviating from an ideal. Our experimental procedure would reflect our interest in that we "standardize" our treatment procedures, and test randomly selected blocks at the different pressures.

A crucial aspect of good experimentation is to designate the important variables as "treatments" and leave the others to be absorbed as "experimental units". One function of statistics is to provide a basis for an objective evaluation of the adequacy of the choice. It is of course obvious that the experimenter never can demonstrate the fact that the characteristics which he believes serve to identify the "treatment" (i.e. the operation he has carried out) do in fact do so.

Later we shall make a distinction between variations in treatment application and variation among experimental units. We note here that

a possible viewpoint would be to regard any departure from the "ideal treatment" as belonging with the "experimental unit" classification. This would, we believe, result in blurring many experimental investigations.

Finally we note that an insistence on identifying every observation by "all possible characteristics" would be equivalent to regarding every event as unique. Patterns of phenomena may not be exact for a given circumstance but the technique im plicit has provided approximate understanding and prediction for wide classes of events.*

2. Types of errors

It is useful to attempt an operational categorization of types of variability and uncertainty which might be of concern in experimental work.

Consider an experiment intended to compare two solvents for their efficiency in extracting some substance from an ore. Since the ore to be used is not entirely homogeneous we could expect different batches of ore will respond differently to the same treatment. Consequently the comparative evaluation of the solvents will be subject to an uncertainty deriving from the heterogeneity of the experimental material. In the same experiment we may intend to work with equal volumes, say 500 ml., of each of the solvents, but owing to the inevitable limitations on technique and equipment the volumes we use will vary somewhat from the intended amounts. Finally, suppose the degree of

*These remarks are intended to refer to general scientific inference and not to statistical inferences resting on probability notions.

extraction is measured by, say, a colorimetric analysis (the reading being a measure of electric current); then we will in general read a value which differs somewhat from the true value due, perhaps, to parallax, vibrations, etc.

The variability due to heterogeneity of experimental units we will refer to as the unit error (or in the language of agronomic experimentation the plot error). The variability due to our inability to reproduce a treatment exactly we will refer to as the treatment error. The variability which comes about in measuring responses we will refer to as the measurement error.

This categorization is neither precise nor exhaustive (for example, in some experiments, there may be "sampling error") but is useful, for these types of errors usually should be regarded differently. The unit errors are, essentially, fixed (unknown) quantities attached to particular experimental units and include all the effects of variable factors which are associated with the units. The treatment and measurement errors are usefully regarded as distinct from unit errors and can often be idealized as independent random variables with means zero.

3. Randomization

An important aspect of most experimental situations is the fact that each experimental unit can be subjected to only one of the treatments of interest. Because of this fact the unit errors contribute to experimental uncertainty. To obtain some control of these unit errors the

device of randomization is used in the statistical design of experiments.

This technique implies, essentially, that random methods of selection and assignment are employed in carrying out the experiment. As a consequence notions of probability and statistical inference may be employed in the objective evaluation of the experimental results. Thus one major function of randomization will be seen to be to "control" unit errors (in a statistical sense).

Another major function of randomization is to permit a wider basis for inference. For example, by randomly selecting a sample of 100 individuals from a group of 10,000 for some test, we can make some inferences (subject to more or less exact measures of uncertainty) about the entire group.

The device of deliberately imposed randomization in application of "ideal treatments" to experimental units usually has no effect on our concept of the treatment and measurement errors. If we feel that these latter types of error can be treated mathematically as random variables, it is because they arise from a multitude of "factors beyond our control",* and not because of the physical act of randomization.

The randomization procedure provides us with an operational (even though somewhat circular) definition of "experimental unit". It is that entity in an experiment to which treatments are assigned at random.

The application of randomization in experimental design is a contribution of R. A. Fisher. The importance of the principle is widely

---

*For some discussion of this point see Cramér (1946).

accepted. There is a considerable literature, however, on controversial views regarding the relative merits of systematic versus randomized experimental layouts. A summary account of this literature is given by Kendall (1946) who lists the appropriate references.

Because of the prominent role of randomization in the work reported on in this document it was felt that it would be of interest to give quotations from various sources on their views of randomization. The remainder of this section is devoted to such quotations. Of necessity, some remarks are "out of context" but we have attempted to avoid misrepresentation of the authors' views.

Fisher (1935a, p. 20):

> "The element in the experimental procedure [of the classic tea-tasting experiment] which contains the essential safeguard is that the two modifications of the test beverage are to be prepared in random order."

Fisher (1935a, p. 29):

> "...it may be said that the simple precaution of randomization will suffice to guarantee the validity of the test of significance, by which the result of the experiment is to be judged."

Fisher (1935a, p. 49):

> "Randomization properly carried out...ensures that the estimates of error will take proper care of all such causes of different growth rates."

Fisher (1935a, p. 71):

> "The validity of our estimate of error...is guaranteed by the

provision that any two plots, not in the same block, shall have the same probability of being treated alike.... The purpose of randomization...is to guarantee the validity of the test of significance."

Anderson and Bancroft (1952, p. 221):

"In field experiments, it is known that the errors in adjacent plots are usually positively correlated. Randomization is used to circumvent much of this difficulty."

Anscombe (1948, p. 1948):

"...the purpose of randomization is that we should be able to ignore the nature of the heterogeneity of the experiment. It neutralizes the incidence of factors whose incidence is independent of the treatments."

Barnard, G. A. (1952, p. 91):

"...the randomization theory assumes less and concludes less, while the Gauss-Markoff theory assumes more and concludes more. It is only with the randomization theory that the conclusions can be said to follow rigorously...from the data and from the data alone."

Bennett and Franklin (1954, p. 491):

"The second requirement of a satisfactory experimental design ...[is]...that the estimates of the effects should be free from bias. The models which were assumed...required that the residual component be randomly distributed with average value zero. In practice this means that any...combination of effects

not under the control of the experimenter should contribute randomly to the various experimental units... it is customary to randomize the contributions of the residual component."

Cochran and Cox (1950, p. 6):

"...we need some means of insuring that a treatment will not be continually favored or handicapped in successive repetitions by some extraneous source of variation, known or unknown. This is done by... randomization...."

"Randomization is somewhat analogous to insurance, in that it is a precaution against disturbances that may or may not occur."

Davies et al. (1954, p. 44):

"...the random allocation of treatments in a suitable statistical design will make it legitimate to analyze the experiment as though the assumption of independence of errors were true."

Greenberg (1951, pp. 310, 313):

"...instead of relying on systematization to balance the effect of order, local control...with randomization, has the effect of removing this disturbance and the errors can be treated as though they were uncorrelated."

"The belief that a systematic arrangement (for a particular situation) would cancel the effect of order was in error. It also fails to fulfill the assumption of randomness implicit in current methods of analysis."

Hald (1952, p. 500):

> "...if the variations caused by the non-experimental factors are not in a state of statistical control, they must be forced to be so, if valid conclusions are to be drawn by means of statistical methods. The remedy is to allot the treatments to the experimental units at random."

Hoel (1954, p. 246):

> "Experiments can often be made valid by applying the principles of randomization and replication."
>
> "...any experiment whose conclusions depend upon these methods [e.g. the analysis of variance] requires randomization."

Hotelling (1947, p. 176):

> "Complete validity...is...dependent on a condition...that...the assignment of personnel and equipment to particular flights must be strictly at random. There is no way of doing this except by the use of the actual mechanism of games of chance, or of their equivalent...."

Kempthorne (1952a, p. 125):

> "The experimenter must decide which of the...causes... [of variation] ...must be controlled experimentally. Those... that he does not control experimentally...he must control...by randomization...only when... the full randomization procedure [is used] is the chain of inductive inference sound."

Mood (1950, p. 318):

> "Statistical inference is impossible in non-randomized experiments."

Pearce (1953, p. 2):

> "...it has been required that each trial should estimate its own accuracy.... The conditions for achieving this... [are] replication and randomization."

Quenouille (1953, p. 5):

> "...it is necessary to allot the treatments to the available material at random if unbiased estimates of both the effect of the treatments and also the reproducibility of the effects are to be obtained."

Yates (1939, p. 441):

> "...by the introduction of randomization into the design...yields can be treated as if their errors were uncorrelated."
>
> "...it is known that the majority of material that the experimenter has to handle does fulfill the required conditions sufficiently nearly, [for the application of least squares analysis] provided that a proper process of randomization is adopted."

C. Models

A considerable part of the present thesis (and in fact all of statistics) is concerned with models, by which we mean a mathematical abstraction supposed to represent the essential relationships in a physical situation. In what follows we attempt an elementary classification of types of models and discuss their relationship. The classes are not mutually exclusive. Recent references which are somewhat relevant are Rafferty (1950), Berkson (1950), Kendall (1951), and

Tukey (1954).

1. Mechanistic models

By a mechanistic model* we mean a functional relationship among quantitative variables representing physical properties in an experimental situation. Traditionally such models derive from the logical analysis of relatively simple mechanical idealizations of actual behavior, coupled of course with "established" empirical facts. An example of this would be an expression giving the velocity of fluid in a pipe as a function of pipe diameter, overall flow and distance from the center, the derivation of such an expression being based in part on concepts of friction in liquid layers and relation of frictional forces to areas of contact and velocities.

A mechanistic model seems to be about the best one can hope for in principle. It should be made explicit that such models, as all others, constitute simply convenient summarizations of knowledge plus a basis for approximate prediction. Several mechanistic models may coexist for a certain phenomenon. A well-known example is the Ideal Gas Law and Van der Waal's Equation of State.

2. Polynomial models

It is well-known from Taylor's Theorem that "well-behaved" functions can be represented approximately, over some range, by a

*This corresponds somewhat, we believe, to what Tukey (1954) called a functional model.

polynomial of sufficiently high degree. This may be thought of as a connecting link between mechanistic models and polynomial models. As an example, specific heat (of gases, say) at constant pressure is known to be dependent on temperature. The functional relation between C_p and T is not known but equations of the form

$$C_p = A_o + A_1 T + A_2 T^2 + A_3 T^3$$

successfully reproduce experimental data over a wide range of temperatures. Such expressions are of course very useful and are employed for prediction or summarization.

As another example, it appears to be common in econometric investigations to "fit" expressions such as

$$y = \beta_o + \beta_1 X_1 + \beta_2 X_2 + \ldots\ldots + \beta_k X_k$$

where the X_i are observable characteristics which it is believed affect y. Such a model may be thought of as approximating a function of k variables by means of a k-dimensional hyperplane, which may be reasonable over a sufficiently small domain of the variables.

The problem of calculating constants so as to get an equation of "best fit" is in the province of statistics.

3. Regression* models

Regression models are essentially a generalization of polynomial models in which the matter of "experimental error" and notions of

*The term regression is used here in the sense of functional relationship with errors rather than in the classical sense involving conditional distributions.

probability are made explicit. The simplest regression model is

$$y = \alpha + \beta X + \epsilon$$

where ϵ is often idealized mathematically as a random variable with certain assumed distributional properties. On the basis of such assumptions estimates of unknown parameters, here α and β, are obtained which under the assumed conditions have desirable properties.

Such procedures appear to have a number of important advantages. They provide an objective method of curve fitting. Even though the statistical assumptions do not ever hold exactly the estimators used will often be reasonable under much weaker conditions. For example, simple least squares will give minimum variance unbiased estimators of α and β if the ϵ are uncorrelated with constant variance. However the estimates will be unbiased if the ϵ are only uncorrelated.

Furthermore, statistical significance test procedures provide an evaluation of whether the proposed model is "good enough". Such tests are two-edged. In the simple case above, if a test of significance of $\beta = 0$ is not significant it may suggest that y and X are not related. Conversely it may suggest that there are important variables whose effect is being summarized in the "error" and the effect of X in determining y is not excessively greater than those unspecified variables.

4. Descriptive models

Descriptive models* are important in experimental statistics for it would appear that it is to this class that most "experimental design" models belong.

Consider a simple example in which we have a women and b men, of unknown weight. Suppose we weigh couples, whose combined weight we represent by y. Then a descriptive model is

$$y_{ij} = \alpha_i + \beta_j$$

where α_i is the "effect" due to the i^{th} woman,

β_j is the "effect" due to the j^{th} man,

y_{ij} is the combined weight of the i^{th} woman and j^{th} man.

For this case, in which we have a good understanding of the relationships, we know that the α_i and β_j represent weights of individuals. But for many situations in which such models are used the physical interpretation of the components is not so immediate, nor for that matter does there appear to be sufficient realization that such explicit interpretation of the components of the model is valuable.

If we refer to women as factor A and men as factor B then it is clear that if we measure combined weights of couples, say in pounds, then factors A and B will be additive, in that (ignoring "errors") every y_{ij} value can be written as the sum of two quantities, one depending on the women only, the other on the man only. We summarize this in the

*Descriptive and definitional models correspond somewhat, we believe, to what Tukey (1954) called tangential models.

statement that for the scale of observation actually used, namely y = pounds, the factors A and B do not interact, or are additive, or all interactions of A and B are zero.

Suppose however instead of measuring weight y in pounds that we had measured $z = y^2$. It would then not be possible to write

$$z_{ij} = \alpha_i' + \beta_j'$$

where α_i' and β_j' are "effects" due to woman i and man j respectively. We would now say that factors A and B interact. This shows that the property of additivity depends not only on the mechanics of the situation but also on the scale of the measurements. To speak of additivity of factors without reference to what is being measured is meaningless.

We formalize the notions somewhat for the case of two factors. An observation y may be said to be additive with respect to two factors A and B, with levels $i = 1, 2, \ldots, a$ and $j = 1, 2, \ldots, b$ respectively, if there exist properties u_i of levels of A and v_j of levels of B and functions $\alpha(u)$ and $\beta(v)$ such that

$$y_{ij} = \alpha_i + \beta_j$$

where $\alpha_i = \alpha(u_i)$

$\beta_j = \beta(v_j)$.

The following statements are of interest. If A and B are additive on the scale y then they are additive on the scale z if and only if z is a linear function of y. If we have a mechanistic model (for two variables, say) and know all the parameters then it will always be possible to find a scale on which the factors (representing the variables) are additive.

However it is not true that if we have a scale on which factors are additive that we can proceed to a mechanistic model, for additivity on a certain scale corresponds to a type of functional structure, which would include an infinity of possible functions.

Additivity is of value from several points of view. First it represents a type of relationship which is easy to think about. Second it gives us an economical and exceedingly useful method of extending our knowledge. (For example the weight of all possible couples could be found from the weights of all men with one woman and of all women with one man, assuming only couples could be weighed.) Third it is a step in the direction of finding a mechanistic model. Fourth, the meaningful interpretation of many statistical experimental designs depend heavily on the (often unpublicized) assumptions (usually with little justification) of additivity.

One of the objectives of Part II of this thesis is the examination of the effects of various types of non-additivity on estimation of effects, estimation of error, estimation of components of variation, and on the interpretation of the analysis of variance. One of the objectives of Part III is to show how descriptive and definitional (see below) statistical models may be used in connection with testing for a certain functional structure relationship of factors.

Models with additivity of factors are not the only possible descriptive models, but do appear to be the most easily handled. For example, in Part III we present some preliminary results for a test of a multiplicative functional structure, which latter amounts to a descriptive

model of the type

$$y_{ij} = (\text{unknown constant}) + (\alpha_i)(\beta_j).$$

5. Definitional models

By a definitional model we mean one which depends only on the experimental situation, procedure and design, and not on specific properties of what is under examination. Such models are of value where methods of random sampling and allocation are used in experimentation; they are useful in focusing our attention on specific relationships (identified as components of the model), either in pinpointing the assumptions which would need to be met if a particular interpretation or inference were to be meaningful, or in indicating where assumptions are unnecessary or unjustified.

In Part II (and also in Part III) we are concerned in much detail with the development of definitional models so we shall not discuss these further here. We note however that descriptive models belong to the class of definitional models in the sense that under particular "assumptions" the definitional models can become descriptive.

6. Probability models

A basic contribution of statistics to science is the introduction of notions of probability into models. The statistical viewpoint is to "live with variation" by making its magnitude estimable and by focusing attention frankly on averages. Some interesting discussion of this point

is given by Tolman (1938).

Either by judicious assumption or by direct action (as in the use of randomization and random sampling) the statistician attempts to summarize extraneous variation as a random variable having partially specified properties.

The most common probability models in experimental statistics are the regression models and the models presented as a basis for the analysis of experimental designs. The study of these is covered by Least Squares Theory and General Linear Hypothesis Theory, (see for example Kempthorne (1952a), Rao (1952), Anderson and Bancroft (1952)).

Various writers have given explicit attention to either the classification or the meaning of probability models which have been used as a basis (or justification) for the analysis of variance. Reference may be made to Neyman et al.(1935), Welch (1937), Eisenhart (1947), Tukey (1949b, 1949c, 1954), Kempthorne (1952a).

D. The Analysis of Variance

The modern technique of the analysis of variance was introduced by R. A. Fisher (1918), and has been used for many purposes in the statistical design and analysis of experiments, as well as in the analysis of nonexperimental data. Roughly and superficially speaking, the analysis of variance procedure consists of the decomposition of the total sum of squares of observations into a varying number of components, which sum to the original total, and some of which components

may be associated with an objective classification of the data. The requirement that the sums of squares "add up" is hardly an essential feature of the technique, though it is in some situations a sufficient condition, under certain strong assumptions of normality of distributions, that the individual sums of squares be independently distributed. This follows from Cochran's Theorem. (See for example Cramér (1951), Kempthorne (1952a)).

1. Uses

The analysis of variance procedure has been applied to make tests of hypotheses; to make tests of statistical significance (using appropriate null hypotheses); to the estimation of measures of variation in "treatment population"; to the estimation of the error of estimates of treatment contrasts; to the estimation of error of estimates of regression coefficients; to give an overall summary of the classification of variation in data; etc. Less definitely, but of perhaps major importance, it has the value of eliciting or providing a framework for the experimenter in designing his experiment, in preliminary analysis of his data and serves as a partial descriptive summary of his data. It would also appear to have considerable potentiality in connection with mechanistic, or functional structure, analysis in experimental situations, though this channel still remains to be explored more fully. (A review of some of the relevant literature in this latter connection is given in Part III.) It is possible that the usual justifications presented for the analysis of variance procedure, emphasizing ad hoc assumptions, has impeded research in this area of functional relationships.

2. The usual assumptions

The usual justification for the analysis of variance procedure in the analysis of randomized experiments, presented in most current statistical texts, is heavily hinged on certain major assumptions involving an assumed linear model all of whose random components are normally and independently distributed. For example, it is usual in the analysis of a randomized block type of experiment to base the analysis on a model

$$y_{ij} = \mu + \beta_i + \tau_j + \epsilon_{ij}$$

where the y_{ij} are the observations, the β_i are "block effects", the τ_j are "treatment effects", and the ϵ_{ij} are "random errors", which are taken to be independently and normally distributed around 0 with (usually) the same unknown variance. As another example, in a two-factor experiment in a completely randomized design the model used might be

$$y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \epsilon_{ijk}$$

where the α_i are the "main effects" of one factor, the β_j are the "main effects" of the other factor, the γ_{ij} are "interactions" of the two factors, and the ϵ_{ijk} are random errors with properties as above.

On the basis of such assumptions, the analysis of variance of randomized experiments may be very elegantly (and suggestively) justified and interpreted as special cases of the statistical mathematical theory

known as General Linear Hypothesis theory. (See for example Kempthorne (1952a), Rao (1952 or Anderson and Bancroft (1952)). From the point of view of estimation the weaker conditions of the Gauss-Markoff Theorem on Least Squares (see for example Rao (1952) or Kempthorne (1952a)) provide justification.

Some indication of the arbitrary character of the assumptions lies in the fact that the models used for experimental designs are invariably "not of full rank", though formally that question is resolved by means of the notion of "estimable functions". (See for example Kempthorne (1952a)).

It is of interest to note the emphasis on randomization in the statistical design of experiments by most writers on experimental design, some of whom base their analyses solely on assumptions such as those indicated above. Some quotations from the literature on the importance of the principle of randomization have been given earlier. However, explicit investigation of the relationship of randomization to the assumptions said to underlie the analysis of variance, and more generally, relevant to the statistical analysis of randomized experiments is not so general. Relevant to this latter matter are references to Fisher (1935, 1936), Neyman et al. (1935), Welch (1937), Pitman (1937), Kempthorne (1952a, 1952b), Wilk (1953a, 1955), and some working papers and reports on the investigations on which the present thesis is based, Wilk (1953b, 1954a, 1954b, 1954c), Wilk and Kempthorne (1954a, 1954b, 1955). As noted elsewhere, a review of some of the relevant literature on randomization has been given by Wilk (1953a).

Some papers which have discussed the classification of models and assumptions underlying the analysis of variance are Fisher (1924), Irwin (1931), Eisenhart (1947), Cochran (1947), Tukey (1949b, 1949c). It should be borne in mind that explicit mathematical statement of models dates in general only from the late '30's, though the models are implicit in previous work.

3. Inadequacies in the usual assumptions

Perhaps the most important deficiency of the usual assumptions is the vagueness concerning the relationship of the model to the experimental situation and design. The emphasis on the various arbitrary assumptions, said to underlie application of the analysis of variance technique, has obscured important questions regarding the use of linear models in the analysis of experiments. It certainly appears odd that it is only necessary to mention that an experiment involves two factors and is of a completely randomized design to be able to write down the model given in the preceding section. If such a model is meaningful it seems clear that it cannot be taken as a causal or mechanistic or functional relationship, since no properties of the experimental situation (e.g. nature of treatments, etc.) are needed to write it down. Thus the important question arises, namely what is the nature and meaning of the linear model?

Another aspect deserving scrutiny is just what portions of the analysis of randomized experiments depend on the strong distributional assumptions usually used. The assumptions of independence and

normality for random errors are always falsified in practice. Does this affect the unbiased estimation of treatment contrasts? Does this affect the estimation of error?

The various independence assumptions which are made by many writers often seem to bear no relationship to the physical situation. Some explicit criticism is made of this by Crump (1951). As part of this same question, there exist what appear to be contradictory viewpoints on the analysis of fairly simple experiments among different writers, each of whom base their analysis on an assumed linear model. For example, the expectations of mean squares and the recommended "error term" for a two-factor mixed-model situation given by Mood (1950), Hald (1952), Mentzer (1953) and Scheffé (1954) differ from that given by Kempthorne (1952a), Anderson and Bancroft (1952), Tukey (1949c), and Bennett and Franklin (1954).

Finally, it would appear to be reasonable to try to bridge the gap between the emphasis on randomization in design and the use of seemingly unconnected, arbitrary, assumptions in the analysis.

E. Relation to Other Work

The object of this Division is to attempt to trace briefly some direct relations of Part II of the present thesis to other work. A review of the literature directly relevant to Part III is given in Part III.

1. General background

The basic contributions of R. A. Fisher to general statistical theory and to the design and analysis of experiments are so widespread as to

challenge referencing. Some items of very direct interest with respect to the present thesis are the technique of the analysis of variance, Fisher (1918, 1924, 1950); the formal introduction of randomization in experimentation, together with discussion on the function of randomization in obtaining unbiased estimates of treatment effects, valid estimates of error, and in randomization tests of significance, Fisher (1926, 1935a, 1936); rules for using the latin square design, Fisher (1935a), Fisher and Yates (1938); discussion of the use of interaction as error, Fisher (1935a); the use of the analysis of variance in testing functional relations, Fisher and Mackenzie (1923).

Of classic importance, too, are the publications of Yates. With particular reference to the present study we list the following contributions: on the function and effect of randomization, Yates (1932, 1936, 1939), Eden and Yates (1933); on the choice of error terms, Yates and Cochran (1938). The great amount of work due to Yates on what are now standard designs is of relevance but is contained in all texts on experimental design.

A review of some of the literature on randomization inference has been given by Wilk (1953a). Contributors to the study of the problem of inference from randomized experiments include Fisher (1926, 1935a, 1936), Yates (1936, 1939), Eden and Yates (1933), Neyman et al. (1936), Welch (1937), Pitman (1937), Yates and Cochran (1938), McCarthy (1937), Anscombe (1948), Grundy and Healy (1950), Kempthorne (1952a, 1952b).

2. "Finite model" analysis

The first explicit mathematical investigation of the role of the "experimental unit", the "true yield", and the function of randomization, with some attention to the definition of a model, appears to be that of Neyman et al (1935). Welch (1937), Pitman (1937) and McCarthy (1937), as well as Neyman et al (1935) utilized these concepts mainly with reference to the question of the validity of the F-test of significance in the analysis of variance of randomized blocks and latin squares.

Kempthorne (1952a) gave detailed attention to the general question of the meaning of linear models in statistical analysis of randomized experiments; showed how explicit models could be defined; showed explicitly the effect of randomization on estimation of treatment effects, estimation of error, and the analysis of variance, under some conditions; discussed the validity of the F-test from the point of view of randomization. The methods and results of Part II of the present thesis are natural generalizations of the "finite model" analyses given by Kempthorne (1952a, 1952b), some extensions of which have been given by Wilk (1953a, 1955).

Kempthorne (1952a, 1952b) considered only the case of a given fixed set of treatments and fixed set of experimental units. His investigations (1952a, 1952b) covered the completely randomized, the randomized block, the latin square, and split-plot designs under restrictive additivity assumptions (explicitly given), and the randomized block design under general conditions.

Wilk (1953a, 1955) studied a generalized design (which included both

the usual completely randomized and randomized block designs as special cases) with no additivity assumptions, but also under the conditions of fixed treatments and units, and found the "bias" in the analysis of variance and in the estimation of error induced by interactions of treatments and experimental units.

Investigations into the problem of specification of error terms for so-called mixed models led to the extension and generalization of these notions to general cases involving the sampling of treatment factors and experimental units. These general investigations were also particularly concerned with the effects of non-additivities on the estimation of treatment effects, on the estimation of error, on the estimation of components of variation, and on the interpretation of the analysis of variance. Some of the results of these investigations have appeared as working papers and reports (Wilk (1953b, 1954a, 1954b, 1954c), Wilk and Kempthorne (1954a, 1954b, 1955)). Part II of the present thesis is the first complete account of the methods and results of these investigations.

3. "Dummy" random variables

Considerable use is made in this thesis of "dummy" random variables in the explicit statement of linear statistical models and in the application of these models to problems of estimation of treatment effects, estima-

tion of error, and interpretation of the analysis of variance. These random variables are of two types, those which are concerned with random selection of a sample of treatments and of experimental units from a population, and those which concern the random allocation of treatments to experimental units according to the restrictions of the experimental design.

The first type are analogous to those introduced by Cornfield (1944). The second type are similar to those used by Kempthorne (1952a). While the two types of dummy random variables have many properties in common they are to be distinguished by virtue of the important difference in their function.

A formal definition of some dummy random variables of the second type, in which they appear as characteristic set functions defined on set-valued random variables, has been given by Wilk (1953a).

4. Other investigations

For a "row by column design" involving the random selection of r rows from R and c columns from C, Tukey (1949c) gave general results (valid for "fixed", "mixed" and "random" as well as intermediate cases) on expected mean squares based on an assumed linear model in which the interaction term was not sampled independently of the main effect terms. He also discussed the generalization of the results for more than two factors.

Results on expected mean squares paralleling those of Tukey (1949c) were obtained independently by Cornfield (1953) using a well-defined

sampling approach, which is applicable to a situation when a population is divisible into, say, an R x C array of primaries with a constant number of sampling units within each primary. A sample is taken from the intersections of r rows and c columns chosen at random with n sampling units for each selected primary.

Bennett and Franklin (1954) have also given general formulae, identical with Tukey's and Cornfield's, for expected mean squares for two and three factors using an assumed linear model with properties similar to that of Tukey (1949c).

The methods of proof used by Tukey (1949c), Cornfield (1953) and Bennett and Franklin (1954) are distinct from one another. Each of these authors considered only the case of equal numbers of observations in the "cells".

With respect to the work reported on in Part II of the present thesis, we should like to note* that the extension of finite model analysis to the case when factors as well as experimental units are sampled was made in July 1953 without knowledge of Tukey's (1949c) investigation, and of course Cornfield's (1953) and Bennett and Franklin's (1954) results. The results (in Part II, Division A, of this thesis) on expected mean squares for a two-factor experiment in a completely randomized design with equal numbers in the cells under the restrictive assumption of unit treatment additivity correspond formally to those given by the authors cited above. However, we note that the results without the assumption of unit

*The remarks of the following three paragraphs are intended to distinguish the present investigations from those carried out by Tukey (1949c), Cornfield (1953) and Bennett and Franklin (1954), but not meant to imply any criticisms or comparative evaluations.

treatment additivity do not correspond.

The bases for the results of Part II are derived statistical models whose properties are consequences of the (possibly random) selection of experimental material and treatments, and of the experimental design.* On the other hand, the above citations give no explicit attention to the experimental design; make no distinction between a two-factor experiment and a randomized block design (which the present investigation does do); pay no attention to the notion of treatment experimental unit interaction (which arises as a natural and integral part of the present investigations); and either do not derive the model used (Tukey (1949c), Bennett and Franklin (1954)), or do not use one (Cornfield (1953)).

Finally we note that the central features of the present investigation are the explicit use of the concept of "experimental unit", "true response" and of the randomization employed in the design, while these aspects are not considered by the authors cited above.

From personal communication we understand that Cornfield and Tukey are collaborating on a joint publication based on their dittoed material cited above.

Hooke (1953, 1954a, 1954b) has given, as part of a more general study of some sampling theory, results and methods of relevance to the problems of Part II of this thesis. He showed the formal identification of Tukey's (1949c) assumed model with the procedure of "sampling from a matrix" (which was essentially what Cornfield (1953) studied), gave a

*By "experimental design" we mean, following Fisher, Yates, Cochran, Kempthorne, etc., the permissible randomization patterns for the application of treatments to experimental units.

special proof of Tukey's (1949c) expected mean squares results, and evaluated variances and covariances of mean squares for the case of sampling from a matrix by rows and by columns.

Kempthorne (1952a) gave a rule for mixed model situations which leads to expected mean squares formally equivalent to those given herein for such situations under restrictive additivity assumptions. Kempthorne (personal communication) does not now regard the derivation given there as logically satisfactory.

Anderson and Bancroft (1952) gave expected mean squares for a randomized block design, when blocks are "random" and treatments "fixed", which are formally equivalent to special cases of results given herein under restrictive additivity assumptions.

II. DERIVED LINEAR MODELS AND THEIR USE IN THE ANALYSIS OF RANDOMIZED EXPERIMENTS

A. The Completely Randomized Design

This part of the thesis deals with the analysis of factorial experiments in a completely randomized design. The essence of the design is that no attempt is made to structure the experimental units, or from another, more accurate, viewpoint no restrictions are imposed in the random assignment of treatments to the available experimental units.

We shall at first describe a rather general experimental situation and design, and develop some basic notation, a population model and a statistical model. Succeeding sections will, however, deal with some special cases in giving detail on the analysis of variance. The most general results are given under Cases 6 and 7. Extension of the results is discussed under Case 8.

1. The experimental situation and design

We shall describe in detail a situation in which the treatment combinations of interest may be classified according to the 'levels' of three 'factors'. This will provide enough generality to indicate the extension of the results. The case of two factors is obtained formally, by considering one factor to have only one level.

Let the factors be denoted by $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$, and suppose that

$\mathcal{A}$ has A levels (let i = 1, 2, . . . , A denote the level of $\mathcal{A}$);

$\mathcal{B}$ has B levels (let j = 1, 2, . . . , B denote the level of $\mathcal{B}$);

$\mathcal{C}$ has C levels (let k = 1, 2, . . . , C denote the level of $\mathcal{C}$).

Suppose further that we have P experimental units, and we wish to study the various treatment combinations with respect to the available experimental units. Let the index m = 1, 2, . . . , P denote the experimental unit in this population of available units.

To investigate this situation the following procedure is to be employed:

(i) Select a levels from A at random (let i* = 1, 2, . . . , a denote the randomly selected levels of $\mathcal{A}$ in order of selection--thus, for example, i* = 1 corresponds to some $i = i_1$);

(ii) Select b levels from B at random (let j* = 1, 2, . . . , b denote the randomly selected levels of $\mathcal{B}$ in order of their selection);

(iii) Select c levels from C at random (let k* = 1, 2, . . . , c denote the randomly selected levels of $\mathcal{C}$ in order of their selection);

(iv) Select p experiment units at random from P(let m* = 1, 2, . . . , p denote the randomly selected units in order of their selection), where

$$p = \sum_{i^*} \sum_{j^*} \sum_{k^*} n_{i^*j^*k^*}, \text{ and all } n_{i^*j^*k^*} \geqslant 1.$$

(v) Apply selected factor levels to selected experimental units at random but so that selected combination (i*j*k*) appears on $n_{i^*j^*k^*}$ experimental units.

We shall make the convention that if, for example, A = a (i.e. all levels of $\mathcal{A}$ are to be studied in the experiment) then i and i* are the same index, so that i* = 1 corresponds to i = 1. This avoids some purely

formal difficulties in the development.

This describes a rather general three-factor experiment. Cases of fixed, mixed and random, as well as "semi-random", models are included as possible special cases. The possibility of equal, proportional or unequal numbers in the subclasses is allowed for. It should be noted that the number of observations associated with a treatment combination depends on the experiment and not on the population of treatments (except in the case of fixed factors). Thus treatment combination (i*j*k*) appears n_{i*j*k*} times but, unless the factors are fixed, the association of (i*j*k*) with values of (ijk) will depend on randomization.

2. The population model

We postulate the existence of a real (unknown) number Y_{ijkm} which represents the 'true' response if unit m is subjected to level i of $\mathcal{A}$, level j of $\mathcal{B}$ and level k of $\mathcal{C}$, and take as our immediate framework of statistical concern the conceptual set $\{Y_{ijkm}\}$, and more particularly certain functions defined on this set. It is implicit that we assume that the true response depends only on the unit and the treatment applied.

To identify the particular quantities of interest we write the algebraic identity (using the usual dot convention to denote means)

$$Y_{ijkm} = Y_{....} + (Y_{i...} - Y_{....}) + (Y_{.j..} - Y_{....}) + (Y_{..k.} - Y_{....})$$
$$+ (Y_{ij..} - Y_{i...} - Y_{.j..} + Y_{....}) + (Y_{i.k.} - Y_{i...} - Y_{..k.} + Y_{....})$$

$$+ (Y_{.jk.} - Y_{.j..} - Y_{..k.} + Y_{....}) + (Y_{ijk.} - Y_{ij..} - Y_{i.k.} + Y_{i...}$$

$$- Y_{.jk.} + Y_{.j..} + Y_{..k.} - Y_{....}) + (Y_{...m} - Y_{....}) + (Y_{ijkm}$$

$$- Y_{ijk.} - Y_{...m} + Y_{....})$$

$$= \mu + a_i + b_j + c_k + (ab)_{ij} + (ac)_{ik} + (bc)_{jk} + (abc)_{ijk} + p_m + q_{ijkm} .$$

Each of the quantities identified above has a physical interpretation:

$\mu = Y_{....}$ is the overall (conceptual) mean response if all levels of all factors were applied to all experimental units.

$a_i = Y_{i...} - Y_{....}$ is the difference between the mean response if all combinations of the i-th level of $\mathcal{A}$ with all levels of $\mathcal{B}$ and $\mathcal{C}$ were applied to all experimental units, and μ. We will call a_i the main effect or simply the effect of the i-th level of $\mathcal{A}$.

Similarly $b_j = Y_{.j..} - Y_{....}$ is the main effect of the j-th level of $\mathcal{B}$, and $c_k = Y_{..k.} - Y_{....}$ is the main effect of the k-th level of $\mathcal{C}$.

$(ab)_{ij} = (Y_{ij..} - Y_{i...} - Y_{.j..} + Y_{....})$ is the difference between the effect of the j-th level of $\mathcal{B}$ in combination with the i-th level of $\mathcal{A}$ and the main effect of the j-th level of $\mathcal{B}$. Thus we call $(ab)_{ij}$ the interaction effect (or simply interaction) of the i-th level of $\mathcal{A}$ and the j-th level of $\mathcal{B}$.

Similarly $(ac)_{ik}$ and $(bc)_{jk}$ are the interaction effects of the i-th level of $\mathcal{A}$ with the k-th level of $\mathcal{C}$, and of the j-th level of $\mathcal{B}$ and the k-th level of $\mathcal{C}$, respectively.

$(abc)_{ijk} = (Y_{ijk.} - Y_{ij..} - Y_{i.k.} + Y_{i...} - Y_{.jk.} + Y_{.j..} + Y_{..k.} - Y_{....})$ measures the difference between the interaction of the j-th level of $\mathcal{B}$ and the k-th level of $\mathcal{C}$ when $\mathcal{A}$ is at the i-th level, and $(bc)_{jk.}$. We

may also, for example, interpret $(abc)_{ijk}$ as the difference between the interaction of the i-th level of $\mathcal{A}$ and the j-th level of $\mathcal{B}$ when $\mathcal{C}$ is at the k-th level, and $(ab)_{ij.}$. In view of the symmetry we can refer to $(abc)_{ijk}$ as the three factor interaction of the i-th level of $\mathcal{A}$, j-th level of $\mathcal{B}$ and k-th level of $\mathcal{C}$.

$p_m = Y_{\ldots m} - Y_{\ldots.}$ measures the difference between the mean response from all combinations of levels of $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$ on unit m, and μ. Thus p_m measures the variability of the experimental units with respect to an "average" of the factors. We call p_m the unit error.

$q_{ijkm} = (Y_{ijkm} - Y_{\ldots m} - Y_{ijk.} + Y_{\ldots.})$ measures the difference between the effect of combination (ijk) on unit m and the effect of combination (ijk) with respect to all units. We therefore refer to q_{ijkm} as the interaction of treatment combination (ijk) with unit m.

It is convenient to introduce here the concept of additivity. The present use of the word is exemplified as follows: two factors will be said to be additive if their interaction is zero. Thus if $(ab)_{ij} = 0$ for all i and j then we say there is additivity of factors $\mathcal{A}$ and $\mathcal{B}$. Similarly if we say that the experimental units are additive with respect to the factor combinations, this means that the interactions q_{ijkm} are zero. These examples should make it clear that the concept of additivity is used to describe the relationships of factors and experimental units for the scale of measurement employed.

It is clear from their definition that

$$0 = \sum_i a_i = \sum_j b_j = \sum_k c_k = \sum_i (ab)_{ij} = \sum_j (ab)_{ij} = \sum_i (ac)_{ik} = \sum_k (ac)_{ik} = \sum_j (bc)_{jk}$$

$$= \sum_k (bc)_{jk} = \sum_i (abc)_{ijk} = \sum_j (abc)_{ijk} = \sum_k (abc)_{ijk} = \sum_m p_m = \sum_{ijk} q_{ijkm}$$

$$= \sum_m q_{ijkm} \cdot$$

We know that if we actually subjected unit m to factor combination (ijk) we would not in general observe the 'true' response, Y_{ijkm}, owing to inevitable errors in applying the factors and in measuring the response. We will assume that if combination (ijk) is applied to unit m we would observe the quantity

$$y_{ijkm} = Y_{ijkm} + \epsilon_{ijkm} ,$$

where the ϵ_{ijkm} can be treated as random variables which are mutually uncorrelated with mean 0 and constant variance σ^2. We shall refer to the ϵ_{ijkm} as the technical errors. The assumption above is not trivial. The largest hurdle is the assumption of additivity of technical error and 'true' response, i.e. no interaction between the unit-factor combination and the technical error. However, the assumption appears to be reasonably valid for a wide class of situations, and is of course quite popular in statistical analyses. (The homogeneity assumption may be weakened somewhat with little complication.)

We can now write for our (conceptual) observable, y_{ijkm}, the following relation

$$y_{ijkm} = \mu + a_i + b_j + c_k + (ab)_{ij} + (ac)_{ik} + (bc)_{jk} + (abc)_{ijk}$$
$$+ p_m + q_{ijkm} + \epsilon_{ijkm} \cdot$$

This relation will be called the population model. No mechanistic notions concerning the relations of the factors are involved in writing

it and the only assumptions used revolve about the technical errors ϵ_{ijkm}. In a general study of this sort we have no recourse but to proceed by assumption with respect to technical errors.

3. The statistical model

We turn now to an analysis of the actual experimental observables. Let $x_{i*j*k*f}$ denote the f-th replicate observation obtained from selected factor combination (i*j*k*), where $f = 1, 2, \ldots, n_{i*j*k*}$. Since each $x_{i*j*k*f}$ comes from some one experimental unit, each (i*j*k*f) corresponds to some value of m, the experimental unit index. The effect of our experimental procedure is that we obtain a random (within the restrictions of the design) sample, the $\{ x_{i*j*k*f} \}$, from the set $\{y_{ijkm}\}$.

It is appropriate to discuss here the function of randomization in this experimental design. Clearly, if we could observe the entire set $\{Y_{ijkm}\}$ we would know everything (empirically) possible about the experimental situation under consideration. Alternately, if we could observe the entire set $\{y_{ijkm}\}$ then the only error involved in our inferences about functions defined on the set $\{Y_{ijkm}\}$ would be the technical errors ϵ_{ijkm}. However, we are able, in general, to observe only a subset of the $\{y_{ijkm}\}$ and hence our inferences will be influenced by additional variabilities. The function of randomization is to attempt to control, in a statistical sense, these additional variabilities, and to enable us to perhaps obtain valid estimates of the uncertainties of inferences.

In order to write an explicit model for the $x_{i^*j^*k^*f}$ in terms of the parameters of interest it is useful to introduce some additional definitions and notation. Let*

$a_i^{i^*}$ = 1 if i* corresponds to i,
= 0 otherwise;

$\beta_j^{j^*}$ = 1 if j* corresponds to j,
= 0 otherwise;

$\gamma_k^{k^*}$ = 1 if k* corresponds to k,
= 0 otherwise;

$\rho_m^{i^*j^*k^*f}$ = 1 if the f-th replicate of selected combination (i*j*k*) corresponds to unit m,
= 0 otherwise.

Because random methods of selection and allocation are employed, the $a_i^{i^*}$, $\beta_j^{j^*}$, $\gamma_k^{k^*}$, $\rho_m^{i^*j^*k^*f}$ are random variables, and from the design of the experiment derive their distributional properties. Thus, for example, the a's, β's, γ's, and ρ's are mutually independent, and

$$P^{**}(a_i^{i^*} = 1) = \frac{1}{A} ;$$

$$P(a_i^{i^*} = 1,\ a_i^{i^{*\prime}} = 1) = 0,\ i^* \neq i^{*\prime} ;$$

$$P(a_i^{i^*} = 1,\ a_{i'}^{i^{*\prime}} = 1) = \frac{1}{A(A-1)} ,\ i \neq i',\ i^* \neq i^{*\prime} ;$$

---

*A more formal definition of related quantities, in which they appear as characteristic set functions defined on set-valued random variables, has been given by Wilk (1953).

**The notation P( ) stands for the probability of occurrence of the event described in the parenthesis.

$$P(\rho_m^{i^*j^*k^*f} = 1) = \frac{1}{P} \ ;$$

$$P(\rho_m^{i^*j^*k^*f} = 1,\ \rho_m^{i^{*\prime}j^{*\prime}k^{*\prime}f^\prime} = 1) = 0,\ (i^*, j^*, k^*, f) \neq (i^{*\prime}, j^{*\prime}, k^{*\prime}, f^\prime);$$

$$P(\rho_m^{i^*j^*k^*f} = 1,\ \rho_{m^\prime}^{i^{*\prime}j^{*\prime}k^{*\prime}f^\prime} = 1) = \frac{1}{P(P-1)},\ m \neq m^\prime,\ (i^*j^*k^*, f) \neq (i^{*\prime}, j^{*\prime}, k^{*\prime}, f^\prime);$$

$$P(a_i^{i^*} a_i^{i^{*\prime}} = 0) = P(a_i^{i^*} a_{i^\prime}^{i^*} = 0) = 1,\ i^* \neq i^{*\prime},\ i \neq i^\prime;$$

etc.

It is thus easy to obtain the expectations of certain functions of these quantities. We list below some representative examples of expectations which are useful in the sequel. Following conventional notation, E(Z) denotes the mathematical expectation of the random variable Z. Then

$$E(a_i^{i^*}) = \frac{1}{A} \ ;$$

$$E(a_i^{i^*} a_{i^\prime}^{i^{*\prime}}) = \frac{1}{A(A-1)},\ i \neq i^\prime,\ i^* \neq i^{*\prime} \ ;$$

$$E(a_i^{i^*} a_i^{i^{*\prime}}) = E(a_i^{i^*} a_{i^\prime}^{i^*}) = 0,\ i \neq i^\prime,\ i^* \neq i^{*\prime} \ ;$$

$$E(\beta_j^{j^*}) = \frac{1}{B} \ ;$$

$$E(\gamma_k^{k^*}) = \frac{1}{C} \ ;$$

$$E(\rho_m^{i^*j^*k^*f}) = \frac{1}{P} \ ;$$

$$E(\rho_m^{i^*j^*k^*f} \rho_{m^\prime}^{i^{*\prime}j^{*\prime}k^{*\prime}f^\prime}) = \frac{1}{P(P-1)},\ (i^*, j^*, k^*, f) \neq (i^{*\prime}, j^{*\prime}, k^{*\prime}, f^\prime),\ m \neq m^\prime;$$

$$E(\rho_m^{i^*j^*k^*f} \rho_{m^\prime}^{i^*j^*k^*f}) = 0,\ m \neq m^\prime \ ,$$

etc.

We can now write an explicit model for the observable random variables $x_{i^*j^*k^*f}$ in terms of the parameters of the population model and

the random variables α_i^{i*} , β_j^{j*} , γ_k^{k*} , $\rho_m^{i*j*k*f}$. Thus

$$x_{i*j*k*f} = \sum_{ijkm} \alpha_i^{i*} \beta_j^{j*} \gamma_k^{k*} \rho_m^{i*j*k*f} (Y_{ijkm} + \epsilon_{ijkm})$$

$$= \mu + \sum_i \alpha_i^{i*} a_i + \sum_j \beta_j^{j*} b_j + \sum_k \gamma_k^{k*} c_k + \sum_{ij} \alpha_i^{i*} \beta_j^{j*} (ab)_{ij}$$

$$+ \sum_{ik} \alpha_i^{i*} \gamma_k^{k*} (ac)_{ik} + \sum_{jk} \beta_j^{j*} \gamma_k^{k*} (bc)_{jk} + \sum_{ijk} \alpha_i^{i*} \beta_j^{j*} \gamma_k^{k*} (abc)_{ijk}$$

$$+ \sum_m \rho_m^{i*j*k*f} p_m + \sum_{ijkm} \alpha_i^{i*} \beta_j^{j*} \gamma_k^{k*} \rho_m^{i*j*k*f} (q_{ijkm} + \epsilon_{ijkm}) .$$

This relationship will be called the statistical model. The random variables in the statistical model are the α_i^{i*} , β_j^{j*} , γ_k^{k*} , $\rho_m^{i*j*k*f}$, which take on the values 0 and 1 with known probability, and the ϵ_{ijkm}. Further, we note that the statistical model derives from the population model by imposing the conditions of the experimental design.

The statistical model can be written in the form

$$x_{i*j*k*f} = \mu + a_{i*}^* + b_{j*}^* + \text{etc.}$$

and the properties of the terms on the right-hand side deduced. This might then be used with homogeneity assumptions to write down a model which is mathematically feasible from the point of view of "exact" distributional theory.

4. Succeeding sections

In the succeeding sections of this Division A we consider special cases, of increasing complexity, of the situation described above. The appropriate specification of the statistical model above is given explicitly for each case considered. In all cases, however, we maintain initially the general relation of, for example, choosing for the experiment a levels of $\mathcal{A}$ from a population of possible A levels of $\mathcal{A}$.

5. Case 1 - one factor, equal numbers, unit-treatment additivity

We first consider the case of one factor under restrictive conditions and assumptions.

Let $\mathcal{A}$ be the factor. We specialize our general description by putting $B = b = C = c = 1$. We can thus write n_{i*j*k*} as n_{i*11} or n_{i*}, and the equal numbers restriction means that $n_{i*} = r$.

Furthermore Y_{ijkm} can be written as Y_{im} and such quantities as b_j, c_k, $(ab)_{ij}$, etc. vanish. The component q_{ijkm} is now simply q_{11lm} and we will assume in this section that units and treatments are additive on the scale employed, that is that the q_{11lm} are all zero.

The statistical model becomes

$$x_{i*f} = \mu + \sum_{i} \alpha_i^{i*} a_i + \sum_{m} \rho_m^{i*f} p_m + \sum_{im} \alpha_i^{i*} \rho_m^{i*f} \epsilon_{im} ,$$

where $f = 1, 2, \ldots, r$ for every $i* = 1, 2, \ldots, a$.

The analysis of variance is given in Table 1. We use the dot convention for means.

Table 1. Analysis of variance. Case 1

Due to	d.f.	Sum of Squares	Mean Square
$\mathcal{A}$	(a-1)	$A' = r \sum_{i*} (x_{i*.} - x_{..})^2$	$A^* = A'/(a-1)$
Residual	a(r-1)	$R' = \sum_{i*f} (x_{i*f} - x_{i*.})^2$	$R^* = R'/a(r-1)$
Total	ar-1	$G' = \sum_{i*f} (x_{i*f} - x_{..})^2$	

Using the statistical model above and the properties of the random variables which derive from the experimental design we can obtain the expectations of the analysis of variance means squares. Table 2 gives the results. The following notational definitions are employed:

$$\sigma_a^2 = \frac{1}{A-1} \sum_i a_i^2 \quad ,$$

$$\sigma_p^2 = \frac{1}{P-1} \sum_m p_m^2 \quad ,$$

$$\sigma^2 = E(\epsilon_{im}^2) \quad .$$

We see that, except for the split up of unit and technical errors, our results are similar to those obtained from the usual assumed linear models.

An unbiased estimate of σ_a^2 is evidently given by

$$\hat{\sigma}_a^2 = \frac{1}{r} (A^* - R^*) \quad .$$

Table 2. Expected mean squares. Case 1

Mean Square	Expected Mean Square
A*	$\sigma^2 + \sigma_p^2 + r\sigma_a^2$
R*	$\sigma^2 + \sigma_p^2$

For the case of a fixed number of levels of $\mathcal{A}$, i.e. A = a, an unbiased estimate of $(a_i - a_{i'})$ is

$$(\hat{a}_i - \hat{a}_{i'}) = (x_{i.} - x_{i'.}), \text{ since}$$

$$(x_{i.} - x_{i'.}) = a_i + \frac{1}{r} \sum_{mf} \rho_m^{ij} (p_m + \epsilon_{im}) - a_{i'} - \frac{1}{r} \sum_{mf} \rho_m^{i'f} (p_m + \epsilon_{i'm}),$$

and

$$E(\sum_m \rho_m^{if} p_m) = \sum_m p_m E(\rho_m^{if}) = \frac{1}{p} \sum_m p_m = 0,$$

$$E(\epsilon_{im}) = 0.$$

The random variable α_i^{i*} does not appear here because, by our convention, for A = a, i* and i are the same index. Hence if i* = 3, say, then

$$P(a_3^3 = 1) = 1$$

$$P(a_i^3 = 0) = 1, \qquad i \neq 3.$$

Thus $\sum_i a_i^3 a_i = a_3$.

The variance of the estimate $(\hat{a}_i - \hat{a}_{i'})$ is

$$E\left[\frac{1}{r} \sum_{mf} \rho_m^{if} (p_m + \epsilon_{im}) - \frac{1}{r} \sum_{mf} \rho_m^{i'f} (p_m + \epsilon_{i'm})\right]^2$$

$$= \frac{2}{r}\sigma^2 + \frac{1}{r^2}\left[E\ (\sum_{mf} \rho_m^{if} p_m)^2 + (\sum_{mf} \rho_m^{i'f} p_m)^2 - 2(\sum_{mf} \rho_m^{if} p_m)(\sum_{mf} \rho_m^{i'f} p_m) \right] .$$

Now, $E(\sum_{mf} \rho_m^{if} p_m)^2 = E(\sum_{mf} \rho_m^{ifj} p_m^2 + \sum_{\substack{m \neq m' \\ f \neq f'}} \rho_m^{if} \rho_{m'}^{if'} p_m p_{m'})$,

since products such as $\rho_m^{if} \rho_m^{if'}$ are 0 with probability 1,

$$= \frac{r}{p} \sum_m p_m^2 + \frac{r(r-1)}{P(P-1)} \sum_{m \neq m'} p_m p_{m'}$$

$$= \frac{r(P-1-r+1)}{P(P-1)} \sum_m p_m^2 .$$

Similarly,

$$E(\sum_{mf} \rho_m^{if} p_m)(\sum_{mf} \rho_m^{i'f} p_m)$$

$$= \frac{2r^2}{P(P-1)} \sum_m p_m^2 .$$

Hence Var $(\hat{a}_i - \hat{a}_{i'}) = \frac{2}{r}(\sigma^2 + \sigma_p^2)$, and is estimated unbiasedly by $\frac{2}{r} R^*$.

For a test of the significance of the data in contradicting the null assumption that $0 = a_1 = a_2 = \text{etc.}$, i.e. that $\sigma_a^2 = 0$, we would choose the criterion A^*/R^*, on the basis of the expectations.

6. Case 2 - one factor, general numbers, additivity

We generalize Case 1 simply to the extent of allowing the possibility of unequal numbers for the various selected levels of the factor. Thus the number of replicates of selected level i^* is $n_{i^*} = ru_{i^*}$, where r is the H. C. F.* of the $\{n_{i^*}\}$. Then for given i^*, the index f has range $1, 2, \ldots, n_{i^*}$.

*H. C. F. is used as an abbreviation for "highest common factor".

We retain the assumption of unit-treatment additivity, and the statistical model is thus the same as Case 1.

The orthogonal analysis of variance (i.e. the individual sums of squares add to the total) is shown in Table 3, and the expected mean squares for this analysis in Table 4. We use the symbol U to denote $\sum_{i*} u_{i*}$, and U* to denote $\sum_{i*} u_{i*}^2/U^2$.

Table 3. Orthogonal analysis of variance. Case 2

Due to	d.f.	Sum of Squares	Mean Square
$\mathcal{A}$	(a-1)	$A' = r \sum_{i*} u_{i*}(x_{i*.} - x_{..})^2$	$A^* = A'/(a-1)$
$\mathcal{R}$	$(r \sum_{i*} u_{i*} - a)$	$R' = \sum_{i*f} (x_{i*f} - x_{i*.})^2$	$R^* = R^*(/(rU-a)$
Total	rU-1	$G' = \sum_{i*f} (x_{i*f} - x_{..})^2$	

Table 4. Expected mean squares. Case 2 (Orthogonal analysis)

Mean Square	Expected Mean Square
A*	$\sigma^2 + \sigma_p^2 + rU \frac{(1-U^*)}{(a-1)} \sigma_a^2$
R*	$\sigma^2 + \sigma_p^2$

Table 5. Non-orthogonal analysis of variance. Case 2

Due to	d.f.	Sum of Squares	Mean Square
$\mathcal{A}$	(a-1)	$A'' = r \sum_{i*} (x_{i*.} - \bar{x}_{..})^2$	$A^{**} = A''/(a-1)$
$\mathcal{R}$	(rU-a)	$R'' = \sum_{i*f} (x_{i*f} - x_{i*.})^2$	$R^{**} = R''/(rU-a)$

It is easily checked that if all n_{i*} are equal then $U = a$ and $U^* = \frac{1}{a}$, whence the coefficient of σ_a^2 in Table 4 becomes the same as it is in Table 2.

Evidently an unbiased estimate of σ_a^2 is given by

$$\hat{\sigma}_a^2 = \frac{(a-1)}{rU(1-U^*)} (A^* - R^*) .$$

For the case of fixed levels of $\mathcal{A}$, i.e. A = a, an unbiased estimate of $(\hat{a}_i - \hat{a}_{i'})$ is

$$(\hat{a}_i - \hat{a}_{i'}) = x_{i.} - x_{i'} .$$

The variance of this estimate is

$$\frac{1}{(n_i + n_{i'})} (\sigma^2 + \sigma_p^2) = \frac{1}{r(u_i + u_{i'})} (\sigma^2 + \sigma_p^2) ,$$

and this is estimated unbiasedly by $\frac{1}{(n_i + n_{i'})} R^*$.

For this situation of unequal numbers one might also consider using a non-orthogonal analysis of variance based on unweighted means, as shown in Table 5. The symbol $\bar{x}_{..}$ denotes the mean of the $x_{i*.}$, i.e. $\bar{x}_{..} = \frac{1}{a} \sum_{i*} x_{i*.}$.

Of course R'' of Table 5 is identical with R' of Table 3. However, A'' of Table 5 is not in general equal to A' of Table 3, unless the n_{i*} are all equal, in which case $u_{i*} = 1$ for all i*.

The expected mean squares for this analysis are given in Table 6.

Table 6. Expected mean squares. Case 2 (Non-orthogonal analysis)

Mean Square	Expected Mean Square
A**	$(\sum_{i*} \frac{1}{u_{i*}})(\sigma^2 + \sigma_p^2) + r\sigma_a^2$
R**	$\sigma^2 + \sigma_p^2$

It may be seen that an unbiased estimate of σ_a^2 is given by

$$\frac{A^{**} - (\sum_{i*} \frac{1}{u_{i*}})R^{**}}{r}$$

We shall have further encounters with similar alternative analyses of variance for several factors with proportional numbers.

7. Case 3 - two factors, equal numbers, unit-treatment additivity

Our attention will be directed, in this section, to a balanced two-factor experiment. Thus we take $\mathcal{A}$ and $\mathcal{B}$ as our factors, put C = c = 1, and omit the k and k* indices in our general development. Components such as $(ac)_{ik'}$ $(bc)_{jk'}$ etc. vanish. We restrict ourself to equal numbers which means that

$$n_{i*j*} = r, \text{ for all } i*j* \quad .$$

We also make the assumption of unit-treatment additivity, i.e. that all $q_{ijm} = 0$.

The statistical model is

$$x_{i*j*f} = \mu + \sum_i \alpha_i^{i*} a_i + \sum_j \beta_j^{j*} b_j + \sum_{ij} \alpha_i^{i*} \beta_j^{j*} (ab)_{ij}$$

$$+ \sum_m \rho_m^{i*j*f} p_m + \sum_{ijm} \alpha_i^{i*} \beta_j^{j*} \rho_m^{i*j*f} \epsilon_{ijm} .$$

The analysis of variance for this situation is given in Table 7. The expected mean squares are given in Table 8. The following definitions are used:

$$\sigma_a^2 = \frac{1}{A-1} \sum_i a_i^2 ,$$

$$\sigma_b^2 = \frac{1}{B-1} \sum_j b_j^2 ,$$

$$\sigma_{ab}^2 = \frac{1}{(A-1)(B-1)} \sum_{ij} (ab)_{ij}^2 ,$$

$$\sigma_p^2 = \frac{1}{P-1} \sum_m p_m^2 ,$$

$$\sigma^2 = E(\epsilon_{ijm}^2) .$$

It will be seen from Table 8 that an unbiased estimate of σ_{ab}^2 is given by

$$\hat{\sigma}_{ab}^2 = \frac{1}{r} (I^*_{AB} - R^*) .$$

Whatever the relation of A and a, an unbiased estimate of σ_b^2 is

$$\hat{\sigma}_b^2 = \frac{1}{ra} \left[B^* - \frac{(A-a)}{A} (I^*_{AB} - R^*) - R^* \right]$$

$$= \frac{1}{ra} \left[B^* - I^*_{AB} + \frac{a}{A} (I^*_{AB} - R^*) \right] .$$

Table 7. Analysis of variance. Case 3

Due to	d.f.	Sum of Squares	Mean Squares
$\mathcal{A}$	(a-1)	$A' = rb \sum_{i*} (x_{i*..} - x_{...})^2$	$A^* = A'/(a-1)$
$\mathcal{B}$	(b-1)	$B' = ra \sum_{j*} (x_{.j*.} - x_{...})^2$	$B^* = B'/(b-1)$
$\mathcal{A} \times \mathcal{B}$	(a-1)(b-1)	$I'_{AB} = r \sum_{i*j*} (x_{i*j*.} - x_{i*..} - x_{.j*.} + x_{...})^2$	$I^*_{AB} = I'_{AB}/(a-1)(b-1)$
Residual	ab(r-1)	$R' = \sum_{i*j*f} (x_{i*j*f} - x_{i*j*.})^2$	$R^* = R'/ab(r-1)$
Total	abr-1	$G' = \sum_{i*j*f} (x_{i*j*f} - x_{...})^2$	

Table 8. Expected mean squares. Case 3

Mean Square	Expected Mean Square
A^*	$\sigma^2 + \sigma_p^2 + r\frac{(B-b)}{B}\sigma_{ab}^2 + rb\sigma_a^2$
B^*	$\sigma^2 + \sigma_p^2 + r\frac{(A-a)}{A}\sigma_{ab}^2 + ra\sigma_b^2$
I^*_{AB}	$\sigma^2 + \sigma_p^2 + r\sigma_{ab}^2$
R^*	$\sigma^2 + \sigma_p^2$

Thus if A = a, i.e. if $\mathcal{A}$ is fixed, then $\hat{\sigma}_b^2$ becomes $\frac{1}{ra}(B^* - R^*)$. Conversely, if A $>>^*$ a, i.e. $\mathcal{A}$ random, then $\hat{\sigma}_b^2$ becomes $\frac{1}{ra}(B^* - I^*_{AB})$.

Similarly an unbiased estimate of σ_a^2 is given by

$$\hat{\sigma}_a^2 = \frac{1}{rb}\left[A^* - I^*_{AB} + \frac{b}{B}(I^*_{AB} - R^*)\right].$$

Suppose now that factor $\mathcal{A}$ is "fixed", i.e. A = a, and we wish to estimate the difference between the responses from levels i and i' of $\mathcal{A}$, averaged overall possible levels of $\mathcal{B}$ and overall experimental units, i.e. $(a_i - a_{i'})$. An unbiased estimate of this is

$$(\hat{a}_i - \hat{a}_{i'}) = (x_{i..} - x_{i'..}),$$

where

$$x_{i..} = \frac{1}{rb}\sum_{j*f}\Big[\mu + a_i + \sum_j \beta_j^{j*} b_j + \sum_j \beta_j^{j*}(ab)_{ij} + \sum_m \rho_m^{ij*f} p_m + \sum_{jm} \beta_j^{j*} \rho_m^{ij*f} \epsilon_{ijm}\Big].$$

This follows because

$$E\left(\sum_j \beta_j^{j*} b_j\right) = \frac{1}{B}\sum_j b_j = 0$$

$$E\left(\sum_j \beta_j^{j*}(ab)_{ij}\right) = \frac{1}{B}\sum_j (ab)_{ij} = 0,$$

etc.

If in fact $\mathcal{B}$ is also fixed, the results are still valid since, with B = b, we take (by convention) j* and j to be the same index.

The variance of $(\hat{a}_i - \hat{a}_{i'})$ is

$$E\Big[\frac{1}{rb}\sum_{j*f}\Big(\sum_j \beta_j^{j*}(ab)_{ij} + \sum_m \rho_m^{ij*f} p_m - \sum_j \beta_j^{j*}(ab)_{i'j} - \sum_m \rho_m^{i'j*f} p_m + \sum_{jm} \beta_j^{j*} \rho_m^{ij*f} \epsilon_{ijm} - \sum_{jm} \beta_j^{j*} \rho_m^{i'j*f} \epsilon_{i'jm}\Big)\Big]^2$$

*We use the symbol >> to stand for 'much greater than'.

$$= \frac{2}{rb}(\sigma^2 + \sigma_p^2) + \frac{1}{b}\frac{(B-b)}{B}\frac{1}{(B-1)}\sum_j\left[(ab^2_{ij} + (ab)^2_{i'j} - 2(ab)_{ij}(ab)_{i'j}\right] .$$

Thus the average variance of differences such as $(\hat{a}_i - \hat{a}_{i'})$ is

$$\frac{2}{rb}(\sigma^2 + \sigma_p^2) + \frac{2}{b}\frac{(B-b)}{B}\sigma^2_{ab} .$$

The selection of an appropriate denominator in setting up a criterion to use in a test of significance will depend upon null hypothesis under scrutiny, as well as on the particular detail of the experiment. For example, if $\mathcal{A}$ is fixed (A = a) while $\mathcal{B}$ is random (B >> b), then we may consider two possible situations. We might wish to have some objective criterion to measure the significance of the data in contradicting either (i) the null assumption that all levels of $\mathcal{A}$ are in fact identical; or (ii) that the effect of all levels of $\mathcal{A}$, averaged over all levels of $\mathcal{B}$, are the same. Formally these null hypotheses are (i) $Y_{ijm} = Y_{i'jm}$, all i, i', j, k; (ii) $\sigma^2_a = 0$. Now the first of these carries the implication that $\sigma^2_a = 0$ and also $\sigma^2_{ab} = 0$.

It would seem reasonable that the criterion A^*/I^*_{AB} be used for hypothesis (ii). However, for hypothesis (i) one might employ $\frac{A^*}{R^*}$, $\frac{A^*}{I^*_{AB}}$ or

$$\frac{(a-1)A^* + (a-1)(b-1)I^*_{AB}}{b(a-1)R^*} .$$

There is some reason to prefer the criterion $\frac{A^*}{R^*}$ in this circumstance. (Further detail for an analogous case is given in Division B of Part II.)

8. Case 4 - two factors, proportional numbers, unit-treatment additivity

We consider now a two factor experiment with "proportional numbers" of observations, under the assumption of additivity of treatments and experimental units. The statistical model is identical with that given under Case 3.

The condition of "proportional numbers" is that

$$n_{i^*j^*} = ru_{i^*}v_{j^*}, \quad i^* = 1, 2, \ldots, a; \quad j^* = 1, 2, \ldots, b;$$

where r is the H. C. F. of the $\{n_{i^*j^*}\}$. An important and often-encountered special case of this is where one factor of classification is unbalanced, as in the following example of numbers of observations:

	$\mathcal{A}1^*$	$\mathcal{A}2^*$	$\mathcal{A}3^*$	
$\mathcal{B}1^*$	4	2	6	12
$\mathcal{B}2^*$	4	2	6	12
	8	4	12	

Here $r = 2$, $u_{1^*} = 2$, $u_{2^*} = 1$, $u_{3^*} = 3$, $v_{1^*} = v_{2^*} = 1$.

It is well-known that the condition of proportional numbers in the subclasses is sufficient so that an orthogonal analysis of variance is possible, where by orthogonal we mean that the individual sums of squares sum to the total sum of squares.

We give in Tables 9 and 10 the orthogonal analysis of variance and corresponding expected mean squares for Case 4. A non-orthogonal analysis based on means might also be used in this case, but we do not

Table 9. Orthogonal analysis of variance. Case 5

Due to	d.f.	Sum of Squares	Mean Square
$\mathcal{A}$	(a-1)	$A' = rV \sum_{i*} u_{i*}(x_{i*..} - x_{...})^2$	$A^* = A'/(a-1)$
$\mathcal{B}$	(b-1)	$B^* = rU \sum_{j*} v_{j*}(x_{.j*.} - x_{...})^2$	$B^* = B'/(b-1)$
$\mathcal{A} \times \mathcal{B}$	(a-1)(b-1)	$I'_{AB} = r \sum_{i*j*} u_{i*}v_{j*}(x_{i*j*.} - x_{i*..} - x_{.j*.} + x_{...})^2$	$I^*_{AB} = I'_{AB}/(a-1)(b-1)$
Residual	(rUV-ab)	$R' = \sum_{i*j*f} (x_{i*j*f} - x_{i*j*.})^2$	$R^* = R'/(rUV-ab)$
Total	rUV-1	$G' = r \sum_{i*j*f} (x_{i*j*f} - x_{...})^2$	

Table 10. Expected mean squares. Case 5 (Orthogonal analysis)

Mean Square	Expected Mean Square
A^*	$\sigma^2 + \sigma_p^2 + rUV \frac{(1-U^*)}{(a-1)} \left[(V^* - \frac{1}{B}) \sigma_{ab}^2 + \sigma_a^2 \right]$
B^*	$\sigma^2 + \sigma_p^2 + rUV \frac{(1-V^*)}{(b-1)} \left[(U^* - \frac{1}{A}) \sigma_{ab}^2 + \sigma_b^2 \right]$
I^*_{AB}	$\sigma^2 + \sigma_p^2 + rUV \frac{(1-U^*)}{(a-1)} \frac{(1-V^*)}{(b-1)} \sigma_{ab}^2$
R^*	$\sigma^2 + \sigma_p^2$

detail it explicitly here. It is however, a special case of a non-orthogonal analysis for unequal numbers given under Case 7.

We use the notation

$$U = \sum_{i*} u_{i*}; \qquad V = \sum_{j*} v_{j*}; \qquad U^* = \sum_{i*} u_{i*}^2/U^2; \qquad V^* = \sum_{j*} v_{j*}^2/V^2 .$$

Note that if $u_{i*} = v_{j*} = 1$ for every i* and j*, then $U = a$, $V = b$, $U^* = \frac{1}{a}$, $V^* = \frac{1}{b}$, and we reproduce the results of Table 8.

Clearly, unbiased estimates of σ^2_{ab}, σ^2_b, and σ^2_a may be obtained from the analysis for this case, but a properly weighted linear combination of the mean squares must be used.

Consider the example of one factor unbalanced given above. For that example, $r = 2$, $a = 3$, $b = 2$, $U = 6$, $V = b = 2$, $U^* = \frac{14}{36}$, $V^* = \frac{1}{2} = \frac{1}{b}$. Then Table 10 becomes:

Mean Square	Expected Mean Square
A*	$\sigma^2 + \sigma^2_p + \frac{11}{3}\frac{(B-2)}{B}\sigma^2_{ab} + \frac{22}{3}\sigma^2_a$
B*	$\sigma^2 + \sigma^2_p + \frac{2}{3}\frac{(7A-18)}{A}\sigma^2_{ab} + 12\,\sigma^2_b$
I^*_{AB}	$\sigma^2 + \sigma^2_p + \frac{11}{3}\sigma^2_{ab}$
R*	$\sigma^2 + \sigma^2_p$

The appropriate combination of mean squares to estimate σ^2_a or σ^2_b will of course depend on the values of A and B.

9. Case 5 - two factors, equal numbers, non-additivity

We now consider the effect of relaxing the assumption of additivity of treatments and experimental units. The example we study in this section is that of a two factor experiment with equal numbers in the subclasses. Thus $n_{i^*j^*} = r$.

The statistical model is

$$x_{i^*j^*f} = \mu + \sum_i \alpha_i^{i^*} a_i + \sum_j \beta_j^{j^*} b_j + \sum_{ij} \alpha_i^{i^*} \beta_j^{j^*} (ab)_{ij} + \sum_m \rho_m^{i^*j^*f} p_m$$

$$+ \sum_{ijk} \alpha_i^{i^*} \beta_j^{j^*} \rho_m^{i^*j^*f} (q_{ijm} + \epsilon_{ijm}) .$$

The analysis of variance is identical with that given in Table 7 for Case 3. The effect of non-additivity of units and treatments on the analysis of variance is partially indicated by the expected mean squares of Table 11.

The following notational definitions are used (in addition to the ones given heretofore):

$$\sigma_q^2 = \frac{1}{AB(P-1)} \sum_{ijm} q_{ijm}^2$$

$$Q_{ap}^2 = \frac{1}{(A-1)(P-1)} \sum_{im} q_{i.m}^2$$

$$Q_{bp}^2 = \frac{1}{(B-1)(P-1)} \sum_{jm} q_{.jm}^2$$

$$Q_{abp}^2 = \frac{1}{(A-1)(B-1)(P-1)} \sum_{ijm} (q_{ijm} - q_{i.m} - q_{.jm})^2 .$$

Table 11. Expected mean squares. Case 5

Mean Square	Expected Mean Square
A*	$\sigma^2 + \sigma^2_p + \sigma^2_q + \frac{r(B-b)}{B}(\sigma^2_{ab} - \frac{1}{P} Q^2_{abp}) + rb(\sigma^2_a - \frac{1}{P} Q^2_{ap})$
B*	$\sigma^2 + \sigma^2_p + \sigma^2_q + \frac{r(A-a)}{A}(\sigma^2_{ab} - \frac{1}{P} Q^2_{abp}) + ra(\sigma^2_b - \frac{1}{P} Q^2_{bp})$
I^*_{AB}	$\sigma^2 + \sigma^2_p + \sigma^2_q + r(\sigma^2_{ab} - \frac{1}{P} Q^2_{abp})$
R*	$\sigma^2 + \sigma^2_p + \sigma^2_q$

It is evident from Table 11 that, in general, non-additivity of units and treatments will introduce biases in the estimation of the components σ^2_{ab}, σ^2_a and σ^2_b. Since each of σ^2_a, σ^2_b, σ^2_{ab}, Q^2_{ap}, Q^2_{bp}, Q^2_{abp} is defined "on the same basis" (i.e. as a sum of squares of quantities divided by the number of linearly independent quantities) the bias will be small if P is large. The smallest value P might have in this situation is rab. Additionally, each of the Q^2 quantities represents a larger order interaction than the σ^2 with which it is associated. For example, σ^2_{ab} reflects the interactions of levels of $\mathcal{A}$ and $\mathcal{B}$, averaged over all experimental units, while Q^2_{abp} reflects the three-way interactions of levels of $\mathcal{A}$ with levels of $\mathcal{B}$ with experimental units. In many situations it will be true that the higher the order of interactions the smaller their magnitude.

The expected mean squares in Table 11 have been presented in such a way as perhaps to emphasize the two extreme special cases, typified by A = a, and A $>>$ a. In view of the interest in fixed factors and random

factors this was thought desirable. However, a more symmetric form exists which permits greater insight into the relationships of the mean squares and is of value from the point of view of extension. The nature of this arrangement is exemplified by

$$E(A^*) = rb(\sigma_a^2 - \frac{1}{B}\sigma_{ab}^2 - \frac{1}{P}Q_{ap}^2 + \frac{1}{BP}Q_{abp}^2) + r(\sigma_{ab}^2 - \frac{1}{P}Q_{abp}^2) + \sigma_q^2 + \sigma_p^2 + \sigma^2$$

$$= rb\,(\sigma_a^2 - \frac{1}{B}\sigma_{ab}^2 - \frac{1}{P}Q_{ap}^2 + \frac{1}{BP}Q_{abp}^2) + r(\sigma_{ab}^2 - \frac{1}{P}Q_{abp}^2)$$

$$+ (\sigma_p^2 - \frac{1}{A}Q_{ap}^2 - \frac{1}{B}Q_{bp}^2 + \frac{1}{AB}Q_{abp}^2) + (Q_{ap}^2 - \frac{1}{B}Q_{abp}^2)$$

$$+ (Q_{bp}^2 - \frac{1}{A}Q_{abp}^2) + Q_{abp}^2 + \sigma^2 \quad .$$

We shall make use of, and give special symbols to, these types of linear combinations of components of variation in the sequel.

The non-additivity of units with treatment combinations does not affect the unbiasedness of linear estimates since

$$E(\sum_{ijk} a_i^{i*}\beta_j^{j*}\rho_m^{i*j*f}q_{ijm}) = 0.$$

However, the average variance of estimates of differences of effects such as

$$(\hat{a}_i - \hat{a}_{i'}) = (x_{i..} - x_{i'..}) ,$$

for the special case of $\mathcal{A}$ fixed (A = a) will now contain a component representing the $\{q_{ijm}\}$. In fact

$$\frac{1}{A(A-1)}\sum_{i\neq i'} Var(\hat{a}_i - \hat{a}_{i'}) = \frac{2}{rb}\left[E(A^*) - rb\sigma_a^2\right] \quad .$$

It would appear then that the effect of non-additivity of units and treatments is (i) to introduce a negative bias into the estimation of the components of variation, (ii) to cause the overestimation of errors of estimates of differences in effects; and (iii) one would guess that it would tend to cause the underestimation of significance levels.

10. Case 6 - three factors, proportional numbers, non-additivity

We now present the case of three factors with no additivity assumptions, but with the number of observations in the subclasses fulfilling the condition that

$$n_{i*j*k*} = r u_{i*} v_{j*} w_{k*} \quad ,$$

where r is the H.C.F. of the n_{i*j*k*} ; i.e. the numbers are proportional. An orthogonal analysis of variance (of weighted means) exists and we shall give the expected mean squares for such an analysis.

The statistical model appropriate for Case 6 is the general one developed above.

We now define in detail certain components which are related to the unit-treatment interactions. We recall that

$$q_{ijkm} = (Y_{ijkm} - Y_{...m} - Y_{ijk.} + Y_{....}) .$$

Hence

$$q_{i..m} = (Y_{i..m} - Y_{...m} - Y_{i...} + Y_{....}) ,$$

$$q_{.j.m} = (Y_{.j.m} - Y_{...m} - Y_{.j..} + Y_{....}) ,$$

$$q_{..km} = (Y_{..km} - Y_{...m} - Y_{..k.} + Y_{....}) ,$$

$$q_{ij.m} = (Y_{ij.m} - Y_{...m} - Y_{ij..} + Y_{....}) ,$$

etc.

and $q_{\ldots m} = q_{ijk.} = 0$.

Let $\sigma_q^2 = \frac{1}{ABC(P-1)} \sum_{ijkm} q_{ijkm}^2$

$Q_{ap}^2 = \frac{1}{(A-1)(P-1)} \sum_{im} q_{i..m}^2$

$Q_{bp}^2 = \frac{1}{(B-1)(P-1)} \sum_{jm} q_{.j.m}^2$

$Q_{cp}^2 = \frac{1}{(C-1)(P-1)} \sum_{km} q_{..km}^2$

$Q_{abp}^2 = \frac{1}{(A-1)(B-1)(P-1)} \sum_{ijm} (q_{ij.m} - q_{i..m} - q_{.j.m})^2$

$Q_{acp}^2 = \frac{1}{(A-1)(C-1)(P-1)} \sum_{ijm} (q_{i.km} - q_{i..m} - q_{..km})^2$

$Q_{bcp}^2 = \frac{1}{(B-1)(C-1)(P-1)} \sum_{jkm} (q_{.jkm} - q_{.j.m} - q_{..km})^2$

$Q_{abcp}^2 = \frac{1}{(A-1)(B-1)(C-1)(P-1)} \sum_{ijkm} (q_{ijkm} - q_{ij.m} - q_{i.km} + q_{i..m} - q_{.jkm} + q_{.j.m} + q_{..km})^2$.

It seems in order to discuss briefly the meaning of these quantities. Each of them may be interpreted as the component of variation due to the interaction of experimental units with factors or with interactions of factors.

Thus σ_q^2 reflects the interaction of experimental units with the treatment combinations. Taking account of the linear restrictions on q_{ijkm} we should have defined σ_q^2 with a divisor of (ABC-1)(P-1). The definition given has been used partly to simplify the formulae, and partly because the quantity σ_q^2 appears in the expectation of the residual mean square. We note that

$$\sigma_q^2 = \frac{(A-1)}{A} Q_{ap}^2 + \frac{(B-1)}{B} Q_{bp}^2 + \frac{(C-1)}{C} Q_{cp}^2 + \frac{(A-1)(B-1)}{AB} Q_{abp}^2$$

$$+ \frac{(A-1)(C-1)}{AC} Q_{acp}^2 + \frac{(B-1)(C-1)}{BC} Q_{bcp}^2 + \frac{(A-1)(B-1)(C-1)}{ABC} Q_{abcp}^2 .$$

The quantity Q_{ap}^2 can also be written as

$$\frac{1}{(A-1)(P-1)} \sum_{im} (Y_{i..m} - Y_{i...} - Y_{...m} + Y_{....})^2 ,$$

which shows explicitly that Q_{ap}^2 represents the interaction of the levels of $\mathcal{A}$ with the experimental units, averaged over all levels of $\mathcal{B}$ and $\mathcal{b}$.

As another example, Q_{abp}^2 may be written as

$$\frac{1}{(A-1)(B-1)(P-1)} \sum_{ijm} (Y_{ij.m} - Y_{i..m} - Y_{ij..} + Y_{i...} - Y_{.j.m} + Y_{.j..} + Y_{...m} - Y_{....})^2 ,$$

which shows that it reflects the interaction of experimental units with the $\mathcal{A} \times \mathcal{B}$ interaction, averaged over all levels of $\mathcal{b}$.

The definition we have used for the Q^2 components is such that their divisors reflect the number of linearly independent quantities of which they are made up. As such they are "on the same basis" as the components σ_a^2, σ_b^2, σ_c^2, σ_{ab}^2, σ_p^2, etc. whose divisors are chosen similarly.

The algebraic structure of the orthogonal analysis of variance sums of squares is a simple extension of that given for Case 4, in Table 9. The expected mean squares for such an analysis, for Case 6, are given in Table 12. Other notational definitions used in Table 12 are as follows:

$$U = \sum_{i*} u_{i*}; \quad V = \sum_{j*} v_{j*}; \quad W = \sum_{k*} w_{k*} ;$$

Table 12. Expectations of mean squares for orthogonal analysis of variance. Case 6

Mean Squares	Expectation
A^*	$\sigma^2 + \sigma^2_q + \sigma^2_p + rUVW\frac{(1-U^*)}{(a-1)}\left\{(V^*-\frac{1}{B})(W^*-\frac{1}{C})\left[\sigma^2_{abc} - \frac{1}{P}Q^2_{abcp}\right]\right.$ $\left. + (W^*-\frac{1}{C})\left[\sigma^2_{ac} - \frac{1}{P}Q^2_{acp}\right] + (V^*-\frac{1}{B})\left[\sigma^2_{ab} - \frac{1}{P}Q^2_{abp}\right] + \left[\sigma^2_a - \frac{1}{P}Q^2_{ap}\right]\right\}$
B^*	$\sigma^2 + \sigma^2_q + \sigma^2_p + rUVW\frac{(1-V^*)}{(b-1)}\left\{(U^*-\frac{1}{A})(W^*-\frac{1}{C})\left[\sigma^2_{abc} - \frac{1}{P}Q^2_{abcp}\right]\right.$ $\left. + (W^*-\frac{1}{C})\left[\sigma^2_{bc} - \frac{1}{P}Q^2_{bcp}\right] + (U^*-\frac{1}{A})\left[\sigma^2_{ab} - \frac{1}{P}Q^2_{abp}\right] + \left[\sigma^2_b - \frac{1}{P}Q^2_{bp}\right]\right\}$
C^*	$\sigma^2 + \sigma^2_q + \sigma^2_p + rUVW\frac{(1-W^*)}{(c-1)}\left\{(U^*-\frac{1}{A})(V^*-\frac{1}{B})\left[\sigma^2_{abc} - \frac{1}{P}Q^2_{abcp}\right]\right.$ $\left. + (V^*-\frac{1}{B})\left[\sigma^2_{bc} - \frac{1}{P}Q^2_{bcp}\right] + (U^*-\frac{1}{A})\left[\sigma^2_{ac} - \frac{1}{P}Q^2_{acp}\right] + \left[\sigma^2_c - \frac{1}{P}Q^2_{cp}\right]\right\}$
I^*_{AB}	$\sigma^2 + \sigma^2_q + \sigma^2_p + rUVW\frac{(1-U^*)(1-V^*)}{(a-1)(b-1)}\left\{(W^*-\frac{1}{C})\left[\sigma^2_{abc} - \frac{1}{P}Q^2_{abcp}\right] + \left[\sigma^2_{ab} - \frac{1}{P}Q^2_{abp}\right]\right\}$
I^*_{AC}	$\sigma^2 + \sigma^2_q + \sigma^2_p + rUVW\frac{(1-U^*)(1-W^*)}{(a-1)(c-1)}\left\{(V^*-\frac{1}{B})\left[\sigma^2_{abc} - \frac{1}{P}Q^2_{abcp}\right] + \left[\sigma^2_{ac} - \frac{1}{P}Q^2_{acp}\right]\right\}$

Table 12 (continued)

Mean Squares	Expectation
I^*_{BC}	$\sigma^2 + \sigma^2_q + \sigma^2_p + rUVW \frac{(1-V^*)(1-W^*)}{(b-1)(c-1)} \left\{(U^* - \frac{1}{A}) \left[\sigma^2_{abc} - \frac{1}{P} Q^2_{abcp}\right] + \left[\sigma^2_{bc} - \frac{1}{P} Q^2_{bcp}\right]\right\}$
I^*_{ABC}	$\sigma^2 + \sigma^2_q + \sigma^2_p + rUVW \frac{(1-U^*)(1-V^*)(1-W^*)}{(a-1)(b-1)(c-1)} \left[\sigma^2_{abc} - \frac{1}{P} Q^2_{abcp}\right]$
R^*	$\sigma^2 + \sigma^2_q + \sigma^2_p$

$$U^* = \frac{\sum_{i^*} u_{i^*}^2}{U^2} ; \quad V^* = \frac{\sum_{j^*} v_{j^*}^2}{V^2} ; \quad W^* = \frac{\sum_{k^*} w_{k^*}^2}{W^2} ;$$

$$\sigma_a^2 = \frac{1}{A-1} \sum_i a_i^2 ; \quad \sigma_b^2 = \frac{1}{B-1} \sum_j b_j^2 ; \quad \sigma_c^2 = \frac{1}{C-1} \sum_k c_k^2 ;$$

$$\sigma_{ab}^2 = \frac{1}{(A-1)(B-1)} \sum_{ij} (ab)_{ij}^2 ; \quad \sigma_{ac}^2 = \frac{1}{(A-1)(C-1)} \sum_{ik} (ac)_{ik}^2 ;$$

$$\sigma_{bc}^2 = \frac{1}{(B-1)(C-1)} \sum_{jk} (bc)_{jk}^2 ; \quad \sigma_{abc}^2 = \frac{1}{(A-1)(B-1)(C-1)} \sum_{ijk} (abc)_{ijk}^2 ;$$

$$\sigma_p^2 = \frac{1}{(P-1)} \sum_m p_m^2 ; \quad \sigma^2 = E(\epsilon_{ijkm}^2) .$$

We note that for the special case of equal numbers $U^* = \frac{1}{a}$, $V^* = \frac{1}{b}$, $W^* = \frac{1}{c}$.

The results of Table 12 indicate, as for Case 5, that in general unbiased estimates of σ_a^2, σ_b^2, σ_c^2, σ_{ab}^2, etc. cannot be obtained from the analysis of variance mean squares if unit-treatment interactions are not negligible. The corresponding statement for the appropriate denominator in a test of significance criterion is complicated by the possible ambiguity with respect to the null hypothesis of concern. But it is apparent that in any test of significance concerning, for example, the levels of $\mathcal{A}$, we cannot find a "denominator" whose expectation will be

$$E(A^* - rUVW \frac{(1-U^*)}{(a-1)} \sigma_a^2) .$$

On the other hand two factors mitigate the preceding comments. Firstly, the quantity confounded with the component of interest enters in the expected mean square with a factor coefficient $\frac{1}{P}$. Thus if the number

of experimental units (i.e., P) is large, then the effect of the combined term will tend to be small. Secondly, each Q^2 quantity represents a higher order interaction term than the component with which it is associated, and is often true that the higher the order of the interaction the smaller it is.

Under the assumption that all $q_{ijkm} = 0$, so-called proper error terms would exist. Table 13 lists error terms for each classification of the design. The bias in using these error terms when unit treatment interactions are not negligible is exemplified by

$$E(A^* - V_A) = rUVW \frac{(1-U^*)}{(a-1)} (\sigma_a^2 - \frac{1}{P} Q_{ap}^2) \quad .$$

The size of unit-treatment interactions depends somewhat independently on two factors, namely the scale of measurement and the heterogeneity of the experimental units. Of course, homogeneous experimental material will mean unit-treatment additivity on any scale.

Apparently the device of randomization is fully effective in allowing unbiased estimates of treatment effects. But essentially unbiased error terms will be obtainable from the analysis of variance, in general, only when the experimental material not too heterogeneous, or when the size, P, of the population of experimental units is large.

11. Case 7 - three factors, general numbers, non-additivity

We consider now the analysis of a three factor experiment, without any additivity assumptions, when no restrictions exist on the numbers n_{i*j*k*}, where n_{i*j*k*} is the number of replicates of the treatment com-

Table 13. Error terms

Classification	Error Term
$\mathcal{A}$	$V_A = R^* + \frac{(b-1)}{(1-V^*)}(V^* - \frac{1}{B})(I^*_{AB} - R^*) + \frac{(c-1)}{(1-W^*)}(W^* - \frac{1}{C})(I^*_{AC} - R^*) - \frac{(b-1)(c-1)}{(1-V^*)(1-W^*)}(V^* - \frac{1}{B})(W^* - \frac{1}{C})(I^*_{ABC} - R^*)$
$\mathcal{B}$	$V_B = R^* + \frac{(a-1)}{(1-U^*)}(U^* - \frac{1}{A})(I^*_{AB} - R^*) + \frac{(c-1)}{(1-W^*)}(W^* - \frac{1}{C})(I^*_{BC} - R^*) - \frac{(a-1)(c-1)}{(1-U^*)(1-W^*)}(U^* - \frac{1}{A})(W^* - \frac{1}{C})(I^*_{ABC} - R^*)$
$\mathcal{C}$	$V_C = R^* + \frac{(a-1)}{(1-U^*)}(U^* - \frac{1}{A})(I^*_{AC} - R^*) + \frac{(b-1)}{(1-V^*)}(V^* - \frac{1}{B})(I^*_{BC} - R^*) - \frac{(a-1)(b-1)}{(1-U^*)(1-V^*)}(U^* - \frac{1}{A})(V^* - \frac{1}{B})(I^*_{ABC} - R^*)$
$\mathcal{A} \times \mathcal{B}$	$V_{AB} = R^* + \frac{(c-1)}{(1-W^*)}(W^* - \frac{1}{C})(I^*_{ABC} - R^*)$
$\mathcal{A} \times \mathcal{C}$	$V_{AC} = R^* + \frac{(b-1)}{(1-V^*)}(V^* - \frac{1}{B})(I^*_{ABC} - R^*)$
$\mathcal{B} \times \mathcal{C}$	$V_{BC} = R^* + \frac{(a-1)}{(1-U^*)}(U^* - \frac{1}{A})(I^*_{ABC} - R^*)$
$\mathcal{A} \times \mathcal{B} \times \mathcal{C}$	$V_{ABC} = R^*$

bination consisting of the i*-th selected level of $\mathcal{A}$, the j*-th selected level of $\mathcal{B}$, and the k*-th selected level of $\mathcal{C}$.

The statistical model appropriate here is the general one developed above.

For a situation involving unequal numbers of observations in the subclasses an orthogonal analysis of variance will not in general exist. The exception to this is the case of proportional numbers, discussed under Case 6.

We give in Table 14 the algebraic structure of an analysis of variance, for this situation, based on the "cell" means. In general, this analysis will not have the property that the individual sums of squares will sum to the so-called total sum of squares. The exception to this statement is when the n_{i*j*k*} are all equal. (Under normal law theory the mean squares of the non-orthogonal analysis would, in general, not be independently distributed.)

In Table 14 we deal with the mean of groups of observations as well as the mean of sets of means. The detailed notation is as follows:

$$\bar{x}_{i*j*..} = \frac{1}{c} \sum_{k*} x_{i*j*k*.} \quad ;$$

$$\bar{x}_{i*.k*.} = \frac{1}{b} \sum_{j*} x_{i*j*k*.} \quad ;$$

$$\bar{x}_{.j*k*.} = \frac{1}{a} \sum_{i*} x_{i*j*k*.} \quad ;$$

$$\bar{x}_{i*...} = \frac{1}{bc} \sum_{j*k*} x_{i*j*k*.} \quad ;$$

$$\bar{x}_{.j*..} = \frac{1}{ac} \sum_{i*k*} x_{i*j*k*.} \quad ;$$

$$\bar{x}_{..k*.} = \frac{1}{ab} \sum_{i*j*} x_{i*j*k*.} \quad ;$$

$$\bar{x}_{....} = \frac{1}{abc} \sum_{i*j*k*} x_{i*j*k*.} \quad ;$$

where $x_{i*j*k*.} = \frac{1}{n_{i*j*k*}} \sum_{f} x_{i*j*k*f}$ is the cell mean. We also write

$N = \sum_{i*j*k*} n_{i*j*k*}$.

It is evident that the sum of squares for the residual of Table 14 is similar in structure to that in the orthogonal analysis of Table 9. The analysis in Table 14 is non-orthogonal (in general) in the sense that the sum of the sums of squares in the table is not equal to

$$\sum_{i*j*k*f} (x_{i*j*k*f} - x_{....})^2 \quad .$$

Of course, for the special case of equal numbers in the cells the analysis given in Tables 7, 9 and 14 become essentially identical. It is useful to note that

$$A'' + B'' + \ldots + I''_{BC} + I''_{ABC} = \sum_{i*j*k*} (x_{i*j*k*.} - \bar{x}_{....})^2 \quad .$$

To aid in the interpretation of this analysis we give in Table 15 the expectations of the mean squares derived from the statistical model. No assumptions are made regarding the unit-treatment interactions $\{q_{ijkm}\}$. We use a double asterisk to denote mean squares for this analysis; e.g. $A^{**} = \frac{A''}{(a-1)}$ = mean square due to $\mathcal{A}$ in this analysis. N^* is used to denote $\sum_{i*j*k*} (\frac{1}{n_{i*j*k*}})$.

The advantage attached to this analysis of variance is the simple structure of the expectations of the mean squares. In fact, if all mean squares, except R^{**}, are adjusted by multiplying by $\frac{abc}{N^*}$, then,

Table 14. The non-orthogonal analysis of variance

Due to	d.f.	Sum of Squares
$\mathcal{A}$	(a-1)	$A'' = bc \sum_{i*} (\bar{x}_{i*..} - \bar{x}_{....})^2$
$\mathcal{B}$	(b-1)	$B'' = ac \sum_{j*} (\bar{x}_{.j*.} - \bar{x}_{....})^2$
$\mathcal{C}$	(c-1)	$C'' = ab \sum_{k*} (\bar{x}_{..k*} - \bar{x}_{....})^2$
$\mathcal{A} \times \mathcal{B}$	(a-1)(b-1)	$I''_{AB} = c \sum_{i*j*} (\bar{x}_{i*j*..} - \bar{x}_{i*...} - \bar{x}_{.j*..} + \bar{x}_{....})^2$
$\mathcal{A} \times \mathcal{C}$	(a-1)(c-1)	$I''_{AC} = b \sum_{i*k*} (\bar{x}_{i*.k*.} - \bar{x}_{i*...} - \bar{x}_{..k*.} + \bar{x}_{....})^2$
$\mathcal{B} \times \mathcal{C}$	(b-1)(c-1)	$I''_{BC} = a \sum_{j*k*} (\bar{x}_{.j*k*.} - \bar{x}_{.j*..} - \bar{x}_{..k*..} + \bar{x}_{....})^2$
$\mathcal{A} \times \mathcal{B} \times \mathcal{C}$	(a-1)(b-1)(c-1)	$I''_{ABC} = \sum_{i*j*k*} (\bar{x}_{i*j*k*.} - \bar{x}_{i*j*..} - \bar{x}_{i*.k*.} + \bar{x}_{i*...} - \bar{x}_{.j*k*.} + \bar{x}_{.j*..} + \bar{x}_{..k*.} - \bar{x}_{....})^2$
$\mathcal{R}$	(N-abc)	$R'' = \sum_{i*j*k*f} (x_{i*j*k*f} - x_{i*j*k*.})^2$

Table 15. Expected mean squares for non-orthogonal analysis of variance. Case 7

Mean Square	Expectation of Mean Square
$A^{**} = \frac{A''}{(a-1)}$	$\frac{N^*}{abc}(\sigma^2 + \sigma^2_q + \sigma^2_p) + \frac{(B-b)}{B}\frac{(C-c)}{C}\left[\sigma^2_{abc} - \frac{1}{P}\Omega^2_{abcp}\right] + \frac{(C-c)}{C} b \left[\sigma^2_{ac} - \frac{1}{P}\Omega^2_{acp}\right]$ $+ \frac{(B-b)}{B} c \left[\sigma^2_{ab} - \frac{1}{P}\Omega^2_{abp}\right] + bc \left[\sigma^2_{a} - \frac{1}{P}\Omega^2_{ap}\right]$
$B^{**} = \frac{B''}{(b-1)}$	$\frac{N^*}{abc}(\sigma^2 + \sigma^2_q + \sigma^2_p) + \frac{(A-a)}{A}\frac{(C-c)}{C}\left[\sigma^2_{abc} - \frac{1}{P}\Omega^2_{abcp}\right] + \frac{(C-c)}{C} a \left[\sigma^2_{bc} - \frac{1}{P}\Omega^2_{bcp}\right]$ $+ \frac{(A-a)}{A} c \left[\sigma^2_{ab} - \frac{1}{P}\Omega^2_{abp}\right] + ac \left[\sigma^2_{b} - \frac{1}{P}\Omega^2_{bp}\right]$
$C^{**} = \frac{C''}{(c-1)}$	$\frac{N^*}{abc}(\sigma^2 + \sigma^2_q + \sigma^2_p) + \frac{(B-b)}{B}\frac{(C-c)}{C}\left[\sigma^2_{abc} - \frac{1}{P}\Omega^2_{abp}\right] + \frac{(B-b)}{B} a \left[\sigma^2_{bc} - \frac{1}{P}\Omega^2_{bcp}\right]$ $+ \frac{(A-a)}{A} b \left[\sigma^2_{ac} - \frac{1}{P}\Omega^2_{acp}\right] + ab \left[\sigma^2_{c} - \frac{1}{P}\Omega^2_{cp}\right]$
$I^{**}_{AB} = \frac{I''_{AB}}{(a-1)(b-1)}$	$\frac{N^*}{abc}(\sigma^2 + \sigma^2_q + \sigma^2_p) + \frac{(C-c)}{C}\left[\sigma^2_{abc} - \frac{1}{P}\Omega^2_{abcp}\right] + c \left[\sigma^2_{ab} - \frac{1}{P}\Omega^2_{abp}\right]$
$I^{**}_{AC} = \frac{I''_{AC}}{(a-1)(c-1)}$	$\frac{N^*}{abc}(\sigma^2 + \sigma^2_q + \sigma^2_p) + \frac{(B-b)}{B}\left[\sigma^2_{abc} - \frac{1}{P}\Omega^2_{abcp}\right] + b \left[\sigma^2_{ac} - \frac{1}{P}\Omega^2_{acp}\right]$

Table 15. (continued)

Mean Square	Expectation of Mean Square
$I^{**}_{BC} = \frac{I''_{BC}}{(b-1)(c-1)}$	$\frac{N^*}{abc}(\sigma^2 + \sigma^2_q + \sigma^2_p) + \frac{(A-a)}{A}\left[\sigma^2_{abc} - \frac{1}{P}\Omega^2_{abcp}\right] + a\left[\sigma^2_{bc} - \frac{1}{P}\Omega^2_{bcp}\right]$
$I^{**}_{ABC} = \frac{I''_{ABC}}{(a-1)(b-1)(c-1)}$	$\frac{N^*}{abc}(\sigma^2 + \sigma^2_q + \sigma^2_p) + (\sigma^2_{abc} - \frac{1}{P}\Omega^2_{abcp})$
$R^{**} = \frac{R''}{(N-abc)}$	$(\sigma^2 + \sigma^2_q + \sigma^2_p)$

speaking rather loosely, this analysis may be interpreted in a similar way to an analysis for a case with equal numbers in the cells.

The discussion given in the preceding section in connection with difficulties when unit-treatment interactions are not negligible apply also to the non-orthogonal analysis. Usually, these interactions will not interfere seriously in the interpretation of this analysis of variance. There does not seem to be any way to circumvent the problems introduced by unit-treatment interactions.

We give in Table 16 what would be "unbiased error terms" for this analysis, for some representative components, if the assumption of negligible unit-treatment interactions were valid.

Table 16. Representative error terms. Case 7

Component	Error Term
σ^2_a	$V'_A = \frac{N^*}{abc} R^{**} + \frac{(B-b)}{B}(I^{**}_{AB} - \frac{N^*}{abc} R^{**})$
σ^2_{ab}	$V'_{AB} = \frac{N^*}{abc} R^{**} + \frac{(C-c)}{C}(I^{**}_{ABC} - \frac{N^*}{abc} R^{**})$
σ^2_{abc}	$V'_{ABC} = \frac{N^*}{abc} R^{**}$

If one prefers, the non-orthogonal analysis described in Tables 14, 15 and 16 may be put in a more familiar form by multiplying all mean squares, except R^{**}, by the factor $\frac{abc}{N^*}$. For the case of equal numbers, i.e., when $n_{i^*j^*k^*} = r$ say, then $\frac{abc}{N^*}$ becomes simply r.

In the present state of knowledge, it appears to be a matter of taste, convenience and opinion as to which analysis is more advantageous when one has a case involving proportional numbers, which is of course a special case of unequal numbers.

It is true that with sufficient distributional assumptions the proportional analysis gives mean squares which are statistically independent. But this property is strictly and uniquely associated with an underlying normal distribution, and so is not to be taken overly seriously. Furthermore, in many cases even this independence property is not directly of use since linear combinations of the mean squares are needed to form a proper error term.

The non-orthogonal analysis has the advantages of wider generality, easier computation, simpler error terms, and more direct connection with the estimation of linear contrasts among treatment effects. Furthermore, speaking very loosely, the non-centrality enters into the mean squares of the non-orthogonal analysis in a simpler and more symmetric fashion than for the orthogonal analysis.

We have seen that the assumption of negligible unit-treatment interactions is not, in general, a critical one in the interpretation of either analysis of variance.

The questions of efficiency of estimation of components of variation and of sensitivity of significance tests, as regards these two analyses, is still open.

12. Case 8 - extension to four or more factors

We discuss briefly in this section the extension of results on expected mean squares to cases of four or more factors in completely randomized designs. In particular we shall consider the case of proportional numbers.

The extension is immediately evident when one considers the results for the three-factor case, given in Table 12, in a more symmetric form. Accordingly, in Table 17 we have detailed the results of Table 12 in this form. The following notational definitions are used in Table 17:

$$\Sigma_a = \sigma_a^2 - \frac{1}{B}\sigma_{ab}^2 - \frac{1}{C}\sigma_{ac}^2 - \frac{1}{P}Q_{ap}^2 + \frac{1}{BC}\sigma_{abc}^2 + \frac{1}{BP}Q_{abp}^2 + \frac{1}{CP}Q_{acp}^2 - \frac{1}{BCP}Q_{abcp}^2$$

Σ_b = Σ_a with b and a, and B and A interchanged.

Σ_c = Σ_a with c and a, and C and A interchanged.

$$\Sigma_{ab} = \sigma_{ab}^2 - \frac{1}{C}\sigma_{abc}^2 - \frac{1}{P}Q_{abp}^2 + \frac{1}{CP}Q_{abcp}^2$$

Σ_{ac} = Σ_{ab} with c and b, and C and B interchanged.

Σ_{bc} = Σ_{ab} with c and a, and C and A interchanged.

$$\Sigma_{abc} = \sigma_{abc}^2 - \frac{1}{P}Q_{abcp}^2$$

$$\Sigma_p = \sigma_p^2 - \frac{1}{A}Q_{ap}^2 - \frac{1}{B}Q_{bp}^2 - \frac{1}{C}Q_{cp}^2 + \frac{1}{AB}Q_{abp}^2 + \frac{1}{AC}Q_{acp}^2 + \frac{1}{BC}Q_{bcp}^2 - \frac{1}{ABC}Q_{abcp}^2$$

$$\Sigma_{ap} = Q_{ap}^2 - \frac{1}{B}Q_{abp}^2 - \frac{1}{C}Q_{acp}^2 + \frac{1}{BC}Q_{abcp}^2$$

$\Sigma_{bp} = \Sigma_{ap}$ with b and a, and B and A interchanged.

$\Sigma_{cp} = \Sigma_{ap}$ with c and a, and C and A interchanged.

$$\Sigma_{abp} = Q^2_{abp} - \frac{1}{C} Q^2_{abcp}$$

$$\Sigma_{acp} = Q^2_{acp} - \frac{1}{B} Q^2_{abcp}$$

$$\Sigma_{bcp} = Q^2_{bcp} - \frac{1}{A} Q^2_{abcp}$$

$$\Sigma_{abcp} = Q^2_{abcp}$$

$$\Sigma_0 = \sigma^2 + \Sigma_{abcp} + \Sigma_{bcp} + \Sigma_{acp} + \Sigma_{abp} + \Sigma_{cp} + \Sigma_{bp} + \Sigma_{ap} + \Sigma_p$$

$$= \sigma^2 + \sigma^2_p + \sigma^2_q .$$

An inverse relationship giving the σ^2 and Q^2 quantities in terms of the Σ's is easily written down. The form of this inverse relationship should be clear by referring to definitions below Table 26, to Table 36, and some discussion in Division E of Part II.

The form of the results given in Table 17 not only makes entirely clear the pattern for extension to more than three factors but also indicates what are, in general, the estimable quantities* in the analysis of variance. It will be evident that an unbiased estimate, based on the analysis of variance mean squares, always exists for each Σ quantity in Table 17. It is of interest that the Σ quantities depend only on the population and not on the sample sizes.

To make explicit the pattern of extensions to more than three factors we give $E(I^*_{AB})$ when we have four factors $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$, $\mathcal{D}$. The notation and definitions implicit should be clear. We use X as analogous to U, V, W, and X* as analogous to U*, V*, W*, with definitions of

*By an estimable quantity we mean one for which an unbiased estimate exists.

Table 17. Symmetric form for the results of Table 12

Mean Squares	Expected Mean Squares
A^*	$rUVW\frac{(1-U^*)}{(a-1)}(\Sigma_a + V^*\Sigma_{ab} + W^*\Sigma_{ac} + V^*W^*\Sigma_{abc}) + \Sigma_0$
B^*	$rUVW\frac{(1-V^*)}{(b-1)}(\Sigma_b + U^*\Sigma_{ab} + W^*\Sigma_{bc} + U^*W^*\Sigma_{abc}) + \Sigma_0$
C^*	$rUVW\frac{(1-W^*)}{(c-1)}(\Sigma_c + U^*\Sigma_{ac} + V^*\Sigma_{bc} + U^*V^*\Sigma_{abc}) + \Sigma_0$
I^*_{AB}	$rUVW\frac{(1-U^*)(1-V^*)}{(a-1)(b-1)}(\Sigma_{ab} + W^*\Sigma_{abc}) + \Sigma_0$
I^*_{AC}	$rUVW\frac{(1-U^*)(1-W^*)}{(a-1)(c-1)}(\Sigma_{ac} + V^*\Sigma_{abc}) + \Sigma_0$
I^*_{BC}	$rUVW\frac{(1-V^*)(1-W^*)}{(b-1)(c-1)}(\Sigma_{bc} + U^*\Sigma_{abc}) + \Sigma_0$
I^*_{ABC}	$rUVW\frac{(1-U^*)(1-V^*)(1-W^*)}{(a-1)(b-1)(c-1)}\Sigma_{abc} + \Sigma_0$
R^*	Σ_0

components of variation as before. Then

$$E(I^*_{AB}) = \Sigma_0 + rUVWX \frac{(1-U^*)(1-V^*)}{(a-1)(c-1)}(W^*X^*\Sigma_{abcd} + W^*\Sigma_{abc} + X^*\Sigma_{abd} + \Sigma_{ab}),$$

(b-1)

where

$$\Sigma_{abcd} = Q^2_{abcd} - \frac{1}{P} Q^2_{abcdp}$$

$$\Sigma_{abc} = \sigma^2_{abc} - \frac{1}{D} \sigma^2_{abcd} - \frac{1}{P} Q^2_{abcp} + \frac{1}{DP} Q^2_{abcdp}$$

$$\Sigma_{ab} = \sigma^2_{ab} - \frac{1}{C} \sigma^2_{abc} - \frac{1}{D} \sigma^2_{abd} - \frac{1}{P} Q^2_{abp} + \frac{1}{CD} \sigma^2_{abcd} + \frac{1}{CP} Q^2_{abcp} + \frac{1}{DP} Q^2_{abdp} - \frac{1}{CDP} Q^2_{abcdp}$$

$$\Sigma_0 = \Sigma_p + \Sigma_{ap} + \Sigma_{bp} + \ldots\ldots + \Sigma_{abcdp} + \sigma^2$$

etc.

Further general discussion on the extension of results for expected mean squares, on a more formal statement of definition of the Σ quantities, and on the reciprocal definition of the σ^2 and Q^2 quantities in terms of the Σ's is given in Division E of this Part.

B. The Randomized Block Design

The essential feature of the randomized block design is that the random allocation of treatments to experimental units is restricted according to a classification of the experimental material. The usual objective of the classification is to try to have the bulk of heterogeneity among units occur between groups, units within groups being more homogeneous.

For example, in testing educational methods, in which test the class is the experimental unit, the various classes might be formed into blocks according to grade, and each method tested in every grade, different methods being randomly assigned to the classes in the grade.

The intention of the blocking in the design is usually to remove the effect of block to block variation from the comparisons of treatments. Sometimes one might also be concerned with the evaluation of block differences with respect to various treatments and with the interaction of blocks with treatments. Thus, in the example, it might be of relevance to compare grades with respect to the various educational methods.

We shall at first describe a general experimental situation and design, develop notation, a population model and a statistical model. Succeeding sections will then deal with special cases and with generalizations.

1. The experimental situation and design

Suppose we have a pool of BP experimental units which are classified into B blocks each containing P experimental units.

Suppose we have T treatments which we wish to examine (compare) with respect to the available population of experimental units.

The experiment is to proceed as follows:

(i) Select t treatments from T at random. ($t \leq T$).

(ii) Select b blocks at random from B. ($b \leq B$).

(iii) From each selected block select at random $p = rt$ experimental units ($r \geq 1$, $p \leq P$).

(iv) Referring to the selected entities, apply treatments to units within blocks at random but so that every treatment appears r times in every block.

Both the experimental situation and the design have features of balance. Thus we assume that the block 'size' (i.e. number of experimental units per block in the population of blocks) is constant, each consisting of P experimental units. In the experimental design, every selected block is represented by the same number, $p = rt$, of experimental units. Further every selected treatment, of which there are t, is replicated the same number, r, of times within each selected block.

This rather general structure contains a number of common special cases. For example, the case of blocks random, treatments fixed is given by $B \gg b$, $T = t$. Most commonly, the randomized block design is employed with $r = 1$. Very often it will be true that $P = p$.

2. The population model

Let $i = 1, 2, \ldots, B$ give the block number in the population of blocks.

Let $j = 1, 2, \ldots, P$ give the unit number within a block, for the

population of experimental units within each block.

Let $k = 1, 2, \ldots, T$ give the treatment number in the population of treatments.

We can conceive of a number which would represent the "true" response, on a given scale, if treatment k were applied to unit j of block i. Let Y_{ijk} be this (conceptual) number. We know that if an experiment were carried out applying treatment k to unit j of block i then we would not in fact observe Y_{ijk} because of inevitable variability in technique of applying the treatment* and/or in measuring the response. Such errors we call technical errors, and we will assume that our (conceptual) observable is a random variable y_{ijk} such that

$$y_{ijk} = Y_{ijk} + \epsilon_{ijk} \quad ,$$

where the ϵ_{ijk} are uncorrelated random variables having mean 0 and variance σ^2 for every i, j, k.

This assumption (it could be relaxed somewhat with little difficulty) on the structure of the technical errors is often reasonable, or at least a reasonable approximation. Since our experimental design does not (usually) involve any control of the technical errors we must proceed by assumption in a general study such as this. For any particular application, this assumption should be reviewed. Of course, if technical errors are small relative to experimental unit heterogeneity then the assumption is of little importance.

Dealing now within our conceptual population of $\{Y_{ijk}\}$, we may write the algebraic identity

*A variation of this point of view is given under Case 6.

$$Y_{ijk} = Y_{...} + (Y_{i..} - Y_{...}) + (Y_{..k} - Y_{...}) + (Y_{i.k} - Y_{i..} - Y_{..k} + Y_{...})$$

$$+ (Y_{ij.} - Y_{i..}) + (Y_{ijk} - Y_{ij.} - Y_{i.k} + Y_{i..})$$

$$= \mu + b_i + t_k + (bt)_{ij} + p_{ij} + n_{ijk} .$$

These quantities have the following physical interpretation:

$\mu = Y_{...}$ is the overall (conceptual) mean response which would be obtained if all treatments were applied to all experimental units in all blocks.

$b_1 = Y_{i..} - \mu$ is the difference between the mean response from all treatments on block i and the overall mean, μ. We call b_i the effect of block i.

Similarly $t_k = Y_{..k} - \mu$ is the effect of treatment k.

$(bt)_{ik} = Y_{i.k} - Y_{i..} - Y_{..k} + Y_{...}$ measures the difference between the effect of treatment k on block i and its effect over all the blocks. We call $(bt)_{ik}$ the interaction of block i and treatment k.

$p_{ij} = Y_{ij.} - Y_{i..}$ measures the difference between the mean response of all treatments on the j-th unit of block i and the mean response of all treatments over all the units of block i. Thus the $\{p_{ij}\}$, $j = 1, 2, \ldots, p$, are a measure of the variation of the units within the i-th block averaged over treatments. We will call p_{ij} the within block additive unit error of the j-th experimental unit of block i.

n_{ijk} measures the difference between the effect of treatment k on unit j of block i and its effect over all of block i. We will refer to the $\{n_{ijk}\}$ as the within block unit-treatment interactions. (A better but more troublesome notation would be $(tp)_{ijk}$.)

We note that, by definition,

$$\sum_i b_i = \sum_k t_k = \sum_i (bt)_{ik} = \sum_k (bt)_{ik} = \sum_j p_{ij} = \sum_j n_{ijk} = \sum_k n_{ijk} = 0 .$$

For our population model we then have

$$y_{ijk} = \mu + b_i + t_k + (bt)_{ik} + p_{ij} + n_{ijk} + \epsilon_{ijk} .$$

3. Relations between population model components

Some consideration of the relationships of the components of the population model is in order. Since the experimental material is generally blocked so that the major sources of variation occur from block to block, it will often be true (depending of course on the scale of measurement) that the block-treatment interactions, the $\{(bt)_{ik}\}$ will be important. It may, of course, be true that the scale of the observations is such that block-treatment additivity exists, i.e. the $(bt)_{ik}$ are negligible; or, alternatively, if the blocks are sufficiently alike the same condition may result. Very often the analysis of randomized block designs is based on either an explicit or implicit assumption of block-treatment additivity. We shall indicate below some of the consequences when this assumption is not fulfilled.

Whatever the magnitude of the $(bt)_{ik}$ it will often be true that experimental units within a block are "sufficiently" alike so that the within block unit treatment interactions, the $\{n_{ijk}\}$, are negligible. This is of course not necessarily true, and the magnitude of the n_{ijk} will depend not only on the degree of variability among units

within blocks, but also on the scale of the observations. In general, the existence of unit treatment interactions has been ignored in the analysis of randomized block designs.

We give below some further discussion of the relations among the components of the population model.

From their definition, the quantities $\{(bt)_{ik}\}$, $\{p_{ij}\}$, and $\{n_{ijk}\}$ are pairwise algebraically independent in groups in that, for example, $p_{ij} = 0$ for all j does not imply mathematically that $n_{ijk} = 0$ for all j. However, in an arrangement of experimental units into blocks, it is generally intended that the major sources of variation in the experimental material should exist between blocks, and that units within blocks should be reasonably alike. Thus from the point of view of the intent of the structure of the design the following statements seem reasonable for many situations:

(i) If the block-treatment interactions $\{(bt)_{ik}\}$ are negligible, so also will be the unit-treatment interactions $\{n_{ijk}\}$.

(ii) If the experimental units within a block are sufficiently alike then generally the bulk of the differences among units within a block will be described by the additive unit errors $\{p_{ij}\}$ and the unit-treatment interactions $\{n_{ijk}\}$ will tend to be negligible.

The following arguments formalize these intuitive notions somewhat. First, focusing attention within a block we note that (a) if the experimental units are identical then $Y_{ijk} = Y_{i.k}$ for all j and k for the given i, and thus the n_{ijk} and also the p_{ij} will be zero; (b) if the treatments are

all identical, but the units are not, then $Y_{ijk} = Y_{ij.}$ for all j and k and thus the n_{ijk} will be zero, but the p_{ij} will not be zero. This suggests that in those situations in which there is a combination of similar treatments and units fairly homogeneous within blocks, the variation of units within blocks will be largely described by the p_{ij}.

A second viewpoint on this matter is based on a Taylor's series approximation argument. We suppose that the conceptual "true" response Y_{ijk} from the k-th treatment on the j-th unit of block i is a function of some property u_{ij} of the unit and some property v_k of the treatment: say

$$Y_{ijk} = f(u_{ij}, v_k) .$$

From the structuring of the experimental units it will often be true that the quantities $u_{i1}, u_{i2}, \ldots, u_{iP}$ are relatively close together (for each i), while the quantities $u_{1.}, u_{2.}, \ldots, u_{B.}$ may be relatively far apart. In general, for similar treatments, $v_1, v_2, \ldots, v_T$ will not be too far apart, but may not be close.

Thus we might approximate $f(u_{ij}, v_k)$ by the first three terms of a Taylor's expansion about the point $(u_{ij}, v_.)$:

$$Y_{ijk} \doteq f(u_{ij}, v_.) + (v_k - v_.) f_2(u_{ij}, v_.) + \frac{(v_k - v_.)^2}{2} f_{22}(u_{ij}, v_.) ,$$

where

$$f_2(u_{ij}, v_.) = \frac{\partial f}{\partial v} \Bigg|_{\substack{u = u_{ij} \\ v = v_.}} ,$$

and

$$f_{22}(u_{ij}, v_{.}) = \frac{\partial^2 f}{\partial v^2}\Bigg|_{\substack{u = u_{ij} \\ v = v_{.}}} .$$

Since u_{ij} is relatively close to $u_{i.}$, it will often be true that

$$f_2(u_{ij}, v_{.}) \doteq f_2(u_{i.}, v_{.}) ,$$

$$f_{22}(u_{ij}, v_{.}) \doteq f_{22}(u_{i.}, v_{.}) ,$$

and $$f(u_{ij}, v_{.}) \doteq f(u_{i.}, v_{.}) + (u_{ij} - u_{i.})\, f_1(u_{i.}, v_{.})$$

where $$f_1(u_{i.}, v_{.}) = \frac{\partial f}{\partial u}\Bigg|_{\substack{u = u_{i.} \\ v = v_{.}}}$$

Then,

$$Y_{ijk} \doteq f(u_{i.}, v_{.}) + (u_{ij} - u_{i.})\, f_1(u_{i.}, v_{.}) + (v_k - v_{.})\, f_2(u_{i.}, v_{.}) + \frac{(v_k - v_{.})^2}{2} f_{22}(u_{i.}, v_{.}) .$$

The next step is to transfer our expansion to the point $(u_{..}, v_{.})$. If the effect of the change from $(u_{i.}, v_{.})$ to $(u_{..}, v_{.})$ is of decreasing importance on successively higher derivatives of the function f, it may be a reasonable approximation that

$$f(u_{i.}, v_{.}) \doteq \bar{f} + (u_{i.} - u_{..})\, \bar{f}_1 + \frac{(u_{i.} - u_{..})^2}{2} \bar{f}_{11} ,$$

$$f_1(u_{i.}, v_{.}) \doteq \bar{f}_1 + (u_{i.} - u_{..})\, \bar{f}_{11} ,$$

$$f_2(u_{i.}, v_{.}) \doteq \bar{f}_2 + (u_{i.} - u_{..})\, \bar{f}_{21} ,$$

and

$$f_{22}(u_{i.}, v_{.}) \doteq \bar{f}_{22} ,$$

where $\bar{f} = f(u_{..}, v_{.})$, $\bar{f}_1 = \frac{\partial f}{\partial u}\Big|_{\substack{u=u_{..} \\ v=v_{.}}}$, $\bar{f}_{21} = \frac{\partial^2 f}{\partial u \partial v}\Big|_{\substack{u=u_{..} \\ v=v_{.}}}$, etc.

Then,
$$\begin{aligned} Y_{ijk} &\doteq \bar{f} + (u_{i.} - u_{..})\bar{f}_1 + \frac{(u_{i.} - u_{..})^2}{2}\bar{f}_{11} \\ &\quad + (u_{ij} - u_{i.})\bar{f}_1 + (u_{ij} - u_{i.})(u_{i.} - u_{..})\bar{f}_{11} \\ &\quad + (v_k - v_{.})\bar{f}_2 + (v_k - v_{.})(u_{i.} - u_{..})\bar{f}_{21} \\ &\quad + \frac{(v_k - v_{..})^2}{2}\bar{f}_{22} \\ &= \bar{f} + (u_{ij} - u_{..})\bar{f}_1 + (u_{i.} - u_{..})(u_{ij} - \tfrac{u_{i.}}{2} - \tfrac{u_{..}}{2})\bar{f}_{11} \\ &\quad + (v_k - v_{.})\bar{f}_2 + (v_k - v_{.})(u_{i.} - u_{..})\bar{f}_{21} + \frac{(v_k - v_{..})^2}{2}\bar{f}_{22} . \end{aligned}$$

If this expansion of the function is reasonable, then one would usually expect the approximation to be better for the various means. Thus

$$Y_{ij.} \doteq \bar{f} + (u_{ij} - u_{..})\bar{f}_1 + (u_{i.} - u_{..})(u_{ij} - \tfrac{u_{i.}}{2} - \tfrac{u_{..}}{2})\bar{f}_{11} + \frac{\bar{f}_{22}}{2T}\sum_k (v_k - v_{.})^2,$$

$$Y_{i..} \doteq \bar{f} + (u_{i.} - u_{..})\bar{f}_1 + \frac{(u_{i.} - u_{..})^2}{2}\bar{f}_{11} + \frac{\bar{f}_{22}}{2T}\sum_k (v_k - v_{.})^2 ,$$

$$Y_{..k} \doteq \bar{f} + \frac{\bar{f}_{11}}{2B}\sum_i (u_{i.} - u_{..})^2 + (v_k - v_{.})\bar{f}_2 + \frac{(v_k - v)^2}{2}\bar{f}_{22} ,$$

$$\begin{aligned} Y_{i.k} &\doteq \bar{f} + (u_{i.} - u_{..})\bar{f}_1 + \frac{(u_{i.} - u_{..})^2}{2}\bar{f}_{11} + (v_k - v_{.})\bar{f}_2 \\ &\quad + (v_k - v_{.})(u_{i.} - u_{..})\bar{f}_{21} + \frac{(v_k - v_{.})^2}{2}\bar{f}_{22} , \end{aligned}$$

$$Y_{...} \doteq \bar{f} + \frac{\bar{f}_{11}}{2B} \sum_i (u_{i.} - u_{..})^2 + \frac{\bar{f}_{22}}{2T} \sum_k (v_k - v_.)^2 \ .$$

Then,

$$b_i \doteq (u_{i.} - u_{..})\,\bar{f}_1 + \left[(u_{i.} - u_{..})^2 - \frac{1}{B} \sum_i (u_{i.} - u_{..})^2\right] \frac{\bar{f}_{11}}{2} \ ,$$

$$t_k \doteq (v_k - v_.)\,\bar{f}_2 + \left[(v_k - v_.)^2 - \frac{1}{T} \sum_k (v_k - v_.)^2\right] \frac{\bar{f}_{22}}{2} \ ,$$

$$(bt)_{ik} \doteq (v_k - v_.)(u_{i.} - u_{..})\,\bar{f}_{21} \ ,$$

$$p_{ij} \doteq (u_{ij} - u_{i.})\,\bar{f}_1 + (u_{i.} - u_{..})(u_{ij} - u_{i.})\,\bar{f}_{11} \ ,$$

$$n_{ijk} \doteq 0.$$

The reasonableness or otherwise of the approximations given above depend of course on the properties of the function f. It is therefore obvious from the preceding argument (as well as from other points of view) that the magnitude of unit-treatment interactions, and also of block-treatment interactions, depends not only on the physical situation (i.e. the mechanics of the reaction, variability of experimental units, etc.) but also on the scale of measurement. For most experiments, it will be true that the magnitude of unit-treatment (and also block-treatment) interactions may be made large by a suitable change of scale in the observed quantity. The converse question is important and interesting but not easily answered.

It can be seen from the steps in the above development that the key assumption leading to the conclusion that $n_{ijk} = 0$ is the one that $f_2(u_{ij}, v_.) = f_2(u_{i.}, v_.)$, which we justified by the presumption that u_{ij} was close to $u_{i.}$. Now a better approximation might be

$$f_2(u_{ij}, v_.) \doteq f_2(u_{i.}, v_.) + (u_{ij} - u_{i.})\, f_{21}(u_{i.}, v_.)$$

$$\doteq \bar{f}_2 + (u_{i.} - u_{..})\, \bar{f}_{21} + (u_{ij} - u_{i.})\, \bar{f}_{21}$$

$$+ (u_{ij} - u_{i.})(u_{i.} - u_{..})\, \bar{f}_{211}$$

$$\doteq \bar{f}_2 + (u_{ij} - u_{..})\, \bar{f}_{21} + (u_{ij} - u_{i.})(u_{i.} - u_{..})\, \bar{f}_{211} \cdot$$

If we can neglect derivatives of order higher than two, then we would have

$$(bt)_{ik} \doteq (v_k - v_.)(u_{i.} - u_{..})\, \bar{f}_{21}$$

$$n_{ijk} \doteq (v_k - v_.)(u_{ij} - u_{i.})\, \bar{f}_{21} \cdot$$

Thus if $(bt)_{ik} = 0$ for all i and k, then $\bar{f}_{21}$ will be negligible and consequently the n_{ijk} will be negligible. However if n_{ijk} is negligible for all i, j and k it may be due to small values of $(u_{ij} - u_{i.})$ and we can conclude nothing about $\bar{f}_{21}$ or the $(bt)_{ik}$.

These considerations suggest that in considering the effect of the n_{ijk} on the interpretation of the analysis of variance it is sufficient to examine those situations in which the block-treatment interactions are not negligible.

4. The statistical model

Whatever the relationship between t, T, b, B, p, P, the design of the experiment is such that, for example, the appearance of treatment k on unit j of block i would preclude the appearance of treatment

k' on unit j of block i, $k \neq k'$. Consequently, our experiment involves essentially the selection for observation of a random (within the restrictions of the design) sample from the set of (conceptual) observable random variables $\{y_{ijk}\}$, $i = 1, 2, \ldots, B$; $j = 1, 2, \ldots, P$; $k = 1, 2, \ldots, T$.

Let $i^* = 1, 2, \ldots, b$ denote the selected blocks, say in order of their selection.

Let $k^* = 1, 2, \ldots, t$ denote the selected treatments in order of their selection.

We shall make the convention that if $B = b$, then i^* and i are to be taken as identical indices. Similarly, if $T = t$, then k^* and k are identical indices.

Let $x_{i^*k^*f}$ represent the observation from the f-th replication of selected treatment k^* in selected block i^*, where f has range $1, 2, \ldots, r$ for every (i^*k^*). We note that for a given i^* each (k^*f) corresponds to some particular j, i.e. any particular replication of a particular selected treatment appears on a particular unit within a given selected block.

It follows from the specification of random selection and allocation that the $x_{i^*k^*f}$ are random variables taking on values in the set of conceptual random variables $\{y_{ijk}\}$. Some of the distributional properties of the $x_{i^*k^*f}$ are determined by the experimental design and procedure and by the properties of the y_{ijk}.

To write an explicit form of the model for the $x_{i^*k^*f}$ in terms of the population parameters it is convenient to introduce some additional definitions and notations.

Let $a_i^{i^*} = 1$ if i^* corresponds to i,

$= 0$ otherwise;

$\beta_k^{k*} = 1$ if k^* corresponds to k,

$= 0$ otherwise;

$\rho_{i*j}^{i*k*f} = 1$ if $(i*k*f)$ corresponds to $(i*j)$,

$= 0$ otherwise.

Because of the random methods of selection and allocation, the quantities $\{a_i^{i*}\}$, $\{\beta_k^{k*}\}$, $\{\rho_{i*j}^{i*k*f}\}$ are random variables. We now specify some of their properties.

$$P(a_i^{i*} = 1) = \frac{1}{B}$$

$$P(a_i^{i*} = 1,\ a_{i'}^{i*'} = 1) = \frac{1}{B(B-1)}, \quad i \neq i', \quad i* \neq i*'$$

$$P(\beta_k^{k*} = 1) = \frac{1}{T}$$

$$P(\beta_k^{k*} = 1,\ \beta_{k'}^{k*'} = 1) = \frac{1}{T(T-1)}, \quad k \neq k', \quad k* \neq k*'$$

$$P(\rho_{i*j}^{k*k*f} = 1) = \frac{1}{P}$$

$$P(\rho_{i*j}^{i*k*f} = 1,\ \rho_{i*j'}^{i*k*'f'} = 1) = \frac{1}{P(P-1)}, \quad j \neq j', \quad (k*f) \neq (k*'f*)$$

etc.

We note that the $\{a_i^{i*}\}$, $\{\beta_k^{k*}\}$, $\{\rho_{i*j}^{i*k*f}\}$ are group-wise statistically independent. Further the sets $\{\rho_{i*j}^{i*k*f}$, all $k*$, f, $j\}$ and $\{\rho_{i*j}^{i*'k*f}$, all $k*f$, $j\}$ are independent for $i* \neq i*'$. This follows from the fact that selected treatments are allotted to selected units in selected blocks at random, independently from block to block.

The following expectations are useful in the sequel:

$$E(a_i^{i*}) = \frac{1}{B}$$

$$E(a_i^{i*}\ a_{i'}^{i*'}) = \frac{1}{B(B-1)}, \quad i \neq i',\ i* \neq i*'$$

$$E(\beta_k^{k*}) = \frac{1}{T}$$

$$E(\beta_k^{k*}\ \beta_{k'}^{k*'}) = \frac{1}{T(T-1)}, \quad k \neq k',\ k* \neq k*'$$

$$E(\rho_{i*j}^{i*k*f}) = \frac{1}{P}$$

$$E(\rho_{i*j}^{i*k*f}\ \rho_{i*j'}^{i*k*'f'}) = \frac{1}{P(P-1)}, \quad j \neq j',\ (k*'f') \neq (k*f)$$

$$E(\rho_{i*j}^{i*k*f}\ \rho_{i*'j'}^{i*'k*'f'}) = \frac{1}{P^2}, \quad i* \neq i*',\ \text{all } k*,\ k*',\ f,\ f',\ j,\ j'.$$

We can now write the following expression, relating the observation x_{i*k*f} to the population parameters and the random variables of the design,

$$x_{i*j(*f} = \sum_{ijk} a_i^{i*}\ \beta_k^{k*}\ \rho_{i*j}^{i*k*f}\ y_{ijk}$$

$$= \mu + \sum_i a_i^{i*} b_i + \sum_k \beta_k^{k*} t_k + \sum_{ik} a_i^{i*}\ \beta_k^{k*} (bt)_{ik}$$

$$+ \sum_{ij} a_i^{i*}\ \rho_{i*j}^{i*k*f}\ p_{ij} + \sum_{ijk} a_i^{i*}\ \beta_k^{k*}\ \rho_{i*j}^{i*k*f} (n_{ijk} + \epsilon_{ijk}) .$$

This expression for x_{i*k*f} we will call the statistical model.

5. Succeeding sections and the analysis of variance

In succeeding sections we deal with special cases under various simplifying conditions, and with generalizations. Under Case 1 we assume that blocks, as well as units within blocks, are additive with respect to treatments on the scale used; i.e. that

$$(bt)_{ik} = n_{ijk} = 0 \text{ for all } i,\ j,\ k.$$

In Case 2 we relax this to the condition that $n_{ijk} = 0$, all i, j, k. In Case 3 we consider the general situation. We maintain the generality of relation between B, b, T, t, P, p and allow $r \geq 1$, but some further discussion of specialized situations is given in the different sections. Case 4 deals with the extension to a factorial structure of treatments; Case 5 considers a hierarchal structure of experimental material; Case 6 involves a hierarchal structure of "treatments".

For each case, the expectations of the analysis of variance mean squares are given. Some discussion is given on problems of estimation, and on the question of choice of criteria for tests of significance.

The analysis of variance is formally the same for Cases 1, 2 and 3, and easily extended to the others, and is detailed in Table 18. The usual dot convention is used to denote means, for example

$$x_{i*..} = \frac{1}{rt} \sum_{k*f} x_{i*k*f} \quad .$$

Table 18. Analysis of variance. General randomized block design

Due to	d.f.	Sum of Squares	Mean Squares
Blocks	(b-1)	$B' = tr \sum_{i*} (x_{i*..} - x_{...})^2$	$B^* = B'/(b-1)$
Treatments	(t-1)	$T' = br \sum_{k*} (x_{.k*.} - x_{...})^2$	$T^* = T'/(t-1)$
Interaction	(b-1)(t-1)	$I' = r \sum_{i*k*} (x_{i*k*.} - x_{i*..} - x_{k*.} + x_{...})^2$	$I^* = I'/(b-1)(t-1)$
Residual	bt(r-1)	$R' = r \sum_{i*k*f} (x_{i*k*f} - x_{i*k*.})^2$	$R^* = R'/bt(r-1)$
Total	(btr-1)	$G' = \sum_{i*k*f} (x_{i*k*f} - x_{...})^2$	

We shall also give here some notational definitions which are employed for Cases 1, 2, and 3; a similar pattern of notation is used for the other cases.

$$\sigma_b^2 = \frac{1}{B-1} \sum_i b_i^2 \quad ;$$

$$\sigma_t^2 = \frac{1}{T-1} \sum_k t_k^2 \quad ;$$

$$\sigma_{bt}^2 = \frac{1}{(B-1)(T-1)} \sum_{ik} (bt)_{ik}^2 \quad ;$$

$$\sigma_p^2 = \frac{1}{B(P-1)} \sum_{ij} p_{ij}^2 \quad ;$$

$$\sigma_n^2 = \frac{1}{BT(P-1)} \sum_{ijk} n_{ijk}^2 \quad ;$$

$$Q_{tp}^2 = \frac{1}{B(T-1)(P-1)} \sum_{ijk} n_{ijk}^2 = \frac{T}{T-1} \sigma_n^2 \quad ;$$

$$\sigma^2 = E(\epsilon_{ijk}^2) \quad .$$

6. Case 1 - block-treatment and unit-treatment additivity

We shall at first consider the case when block-treatment and unit-treatment interactions are negligible. Formally, this is the assumption that

$$(bt)_{ik} = n_{ijk} = 0, \qquad \text{all } i, j, k.$$

The statistical model then becomes

$$x_{i*k*f} = \mu + \sum_i \alpha_i^{i*} b_i + \sum_k \beta_k^{k*} t_k + \sum_{ij} \alpha_i^{i*} \rho_{i*j}^{i*k*f} p_{ij} + \sum_{ijk} \alpha_i^{i*} \beta_k^{k*} \rho_{i*j}^{i*k*f} \epsilon_{ijk} \quad .$$

This could be written in the form

$$x_{i*k*f} = \mu + b_{i*}^* + t_{k*}^* + e_{i*k*f}^* \quad ,$$

with the obvious correspondence between the two forms. This latter form is, of course, reminiscent of the models usually employed for analysis of randomized block experiments.

Whether a model such as this is valid and meaningful depends on the particular circumstances. It should be noted however that if a model, such as the one immediately above, has meaning on the scale of x then a model of that type will not, in general, be valid for a function of x such as, for example, x^2 or log x.

We give in Table 19 the expectations of the analysis of variance mean squares (given in Table 18) using the model of this Case 1.

Table 19. Expected mean squares. Case 1

Mean Square	Expectation of Mean Square
B*	$\sigma^2 + \frac{(P-p)}{P}\sigma_p^2 + tr\sigma_b^2$
T*	$\sigma^2 + \sigma_p^2 + br\sigma_t^2$
I*	$\sigma^2 + \sigma_p^2$
R*	$\sigma^2 + \sigma_p^2$

It is evident that under the conditions of Case 1 I* and R* each estimate the same quantity. Where it was felt that the conditions of Case 1 attained the design would usually specify r = 1, and then R* would not exist.

Evidently an unbiased estimate of σ_t^2 is given under these conditions, by

$$\frac{1}{br}\left[T^* - \frac{(b-1)(t-1)\,I^* + bt\,(r-1)R^*}{(btr-b-t+1)}\right]$$

whatever the values of B, b, T, t, P, and r.

The lack of symmetry with respect to the estimation of σ_b^2 and σ_t^2 is evident from Table 19. Only if $P >> p$ (i.e. the block size is much larger than the number of units actually used) can we obtain unbiased estimates of σ_b^2 .

The unbiased estimation of the errors of contrasts among treatment effects, the t_k, if treatments are fixed ($T = t$), is always possible under the conditions of Case 1. Of course, unbiased estimates of contrasts of the t_k will exist under more general conditions.

Thus, if $\hat{t}_k = (x_{.k.} - x_{...})$, then an unbiased estimate of $\sum_k c_k t_k$ is given by $\sum_k c_k \hat{t}_k$, as is easily checked from the statistical model.

The variance of any estimated contrast $\sum_k c_k \hat{t}_k$, with $\sum_k c_k = 0$, is

$$\frac{\left(\sum_k c_k^2\right)}{rb}\,(\sigma^2 + \sigma_p^2) .$$

7. Case 2 - unit-treatment additivity

We relax the assumptions of Case 1 to take account of interactions of blocks and treatments, but maintain the assumption that units and treatments are additive within blocks. In view of the objective of blocking, which is to obtain reasonably homogeneous units within a block, this assumption may not be unreasonable for many different scales of

measurement.

The statistical model now becomes

$$x_{i*k*f} = \mu + \sum_i a_i^{i*} b_i + \sum_k \beta_k^{k*} t_k + \sum_{ik} a_i^{i*} \beta_k^{k*} (bt)_{ik} + \sum_{ij} a_i^{i*} \rho_{i*j}^{i*k*f} p_{ij}$$

$$+ \sum_{ijk} a_i^{i*} \beta_k^{k*} \rho_{i*j}^{i*k*f} e_{ijk} \quad .$$

We give in Table 20 the expectation of mean squares for the conditions of Case 2, obtained using the statistical model above.

Table 20. Expected mean squares. Case 2

M. S.	Expectation of Mean Square
B*	$\sigma^2 + \frac{(P-p)}{p} \sigma_p^2 + \frac{(T-t)}{T} r\sigma_{bt}^2 + tr\sigma_b^2$
T*	$\sigma^2 + \sigma_p^2 + \frac{(B-b)}{B} r\sigma_{bt}^2 + br\sigma_t^2$
I*	$\sigma^2 + \sigma_p^2 + r\sigma_{bt}^2$
R*	$\sigma^2 + \sigma_p^2$

We will summarize here the experimental conditions and assumptions underlying Table 20 of expected mean squares. These are:

(1) We randomly selected a sample of size b from a population of B blocks of experimental units.

(2) Each block contains P experimental units.

(3) We randomly select a sample of size p = tr from each of the selected b blocks.

(4) We randomly select t treatments from a population of T treatments.

(5) We randomly allocate selected treatments to selected blocks of selected experimental units so that each treatment appears r times in each block.

(6) Treatments are additive with respect to experimental units within blocks; i.e. $n_{ijk} = 0$ all i, j, k.

(7) The technical errors, associated with placing any treatment on any experimental unit and with measuring the response, are additive with respect to the "true" response, and may be treated as random variables which are mutually uncorrelated with mean 0 and constant variance.

Of these, (1) to (5) inclusive must be satisfied by the experimental conditions and design; (6) may be reasonable in many circumstances. For a general study such as this we have no recourse but assumption with respect to technical errors.

It is of interest to note that the function of random allocation of treatments to experimental units is to "control", in a statistical sense, the variation among experimental units within a block. Our notions about the technical error, due to variations in technique, usually are unaffected by the randomization procedure. It is of course obvious that the only basis for making a statistical inference over some wider grouping, say of blocks, than is used in the experiment, is that our selection procedure allows us to treat the blocks actually used as a random sample from the larger aggregate.

Under the condition that blocks and treatments are not additive it is evident that the mean square R* will, in general, be useful and perhaps essential in the interpretation and evaluation of the experiment. Thus one would want if possible to have $r > 1$. (With unit-treatment additivity it would not be necessary to replicate all selected treatments within all selected blocks to obtain an unbiased estimate of $\sigma^2 + \sigma_p^2$, but for simplicity we have considered the balanced procedure).

For conciseness we will use the phrase error term in a sense exemplified by: the error term for σ_t^2 is a quantity V_t such that $E(V_t) = E(T^* - br\sigma_t^2)$. The relation of error terms to estimation of components of variation, to criteria for significance tests, and to the estimation of errors of contrast estimates will be evident.

We consider two subcases.

(a) $r > 1$. (Hence R* exists).

(i) The error term for σ_{bt}^2 is $V_{bt} = R^*$.

(ii) The error term for σ_t^2 is $V_t = I^* - \frac{b}{B}(I^* - R^*)$; when $B = b$ (blocks "fixed") $V_t = R^*$, and when $B >> b$ (blocks "random") $V_t = I^*$.

(iii) If $P \overset{*}{\leftrightarrow} p$, there is no error term for σ_b^2.

(iv) If $P >> p$, then the error term for σ_b^2 is $V_b = I^* - \frac{t}{T}(I^* - R^*)$; thus if $T = t$ (treatments fixed), $V_b = R^*$, and when $T >> t$ (treatments random), $V_b = I^*$.

(v) If technical errors are negligible, i.e. $\sigma^2 = 0$, then the error term for blocks is $(I - \frac{t}{T})I^* - (\frac{p}{P} - \frac{t}{T})R^*$; thus if $T = t$ and

---

*The symbol $\leftrightarrow$ is used here to denote "of the same order of magnitude".

P = p, then block effects for the blocks actually used are known without error.

(b) r = 1. (Hence R* does not exist).

(i) There is no error term for σ^2_{bt} .

(ii) If B $\leftrightarrow$ b, there is no error term for σ^2_t .

(iii) If B >> b, the error term for σ^2_t is I*.

(iv) If P$\leftrightarrow$p and/or T$\leftrightarrow$t there is no error term for σ^2_b.

(v) If P >> p and T >> t then the error term for σ^2_b is I*.

It is apparent that if block-treatment interactions are not negligible then it will usually be useful to have within-block replication of treatments.

When treatments are "fixed" (T = t) one will often be interested in estimating certain contrasts among the treatment effects, in particular differences. Under the conditions of Case 2, an unbiased estimate of $(t_k - t_{k'})$ is given by

$$(\hat{t}_k - \hat{t}_{k'}) = (x_{.k.} - x_{.k'.}) .$$

(The preceding statement in no way depends on the value of r). The variance of this estimate is

$$V(\hat{t}_k - \hat{t}_{k'}) = \frac{2}{br}(\sigma^2 + \sigma^2_p) + \frac{1}{b}\frac{(B-b)}{B}\frac{1}{(B-1)}\sum_i \left[(bt)^2_{ik} + (bt)^2_{ik'} - 2(bt)_{ik}(bt)_{ik'}\right].$$

Thus the average variance of such estimates of differences of treatment effects is

$$\frac{1}{T(T-1)}\sum_{k\neq k'} V(\hat{t}_k - \hat{t}_{k'}) = \frac{2}{br}(\sigma^2 + \sigma^2_p) + \frac{2}{b}\frac{(B-b)}{B}\sigma^2_{ab} .$$

An unbiased estimate of this average variance of estimates of differences of treatment effects is given by

$$\frac{2}{br}\left[R^* + \frac{(B-b)}{B}(I^* - R^*)\right].$$

For general values of B and b, it will be seen that R* is needed to obtain unbiased estimates of the error. In the particular case when $B >> b$, then the unbiased estimate of the average variance becomes $\frac{2}{br} I^*$.

Some discussion on the choice of a criterion for significance tests, given under Case 3 below, is directly relevant also to the conditions of Case 2.

8. Case 3 - general conditions

The results of this section are independent of any assumptions of additivity of treatments with blocks or with units within blocks. The appropriate statistical model for this case is the general one we have developed above.

The analysis of variance is the one given in Table 18. The expectations of mean squares under the general conditions of Case 3 are given in Table 21. Notational definitions used are those given above.

We have presented the results in Table 21 in such a way as to emphasize the two extremes of "fixed" and "random", exemplified for blocks by the conditions $B = b$ and $B >> b$, respectively. The same results are given also in Table 22 in what is perhaps a more instructive form.

Table 21. Expected mean squares. Case 3

Mean Square	Expected Mean Square
B*	$\sigma^2 + \sigma_n^2 + \frac{(P-p)}{P}\sigma_p^2 + \frac{(T-t)}{T}r\left[\sigma_{bt}^2 - \frac{1}{P}\Omega_{tp}^2\right] + tr\sigma_b^2$
T*	$\sigma^2 + \sigma_n^2 + \sigma_p^2 + \frac{(B-b)}{B}r\sigma_{bt}^2 + r\left[b\sigma_t^2 - \frac{1}{P}\Omega_{tp}^2\right]$
I*	$\sigma^2 + \sigma_n^2 + \sigma_p^2 + r\left[\sigma_{bt}^2 - \frac{1}{P}\Omega_{tp}^2\right]$
R*	$\sigma^2 + \sigma_n^2 + \sigma_p^2$

Table 22. Expected mean squares. Case 3 (Alternate form)

Mean Square	Expected Mean Square
B*	$rt\left[\sigma_b^2 - \frac{1}{T}\sigma_{bt}^2 - \frac{1}{P}\sigma_p^2 + \frac{1}{PT}\Omega_{tp}^2\right] + r\left[\sigma_{bt}^2 - \frac{1}{P}\Omega_{tp}^2\right] + \left[\sigma_p^2 - \frac{1}{T}\Omega_{tp}^2\right] + \Omega_{tp}^2 + \sigma^2$
T*	$rb\left[\sigma_t^2 - \frac{1}{B}\sigma_{bt}^2\right] + r\left[\sigma_{bt}^2 - \frac{1}{P}\Omega_{tp}^2\right] + \left[\sigma_p^2 - \frac{1}{T}\Omega_{tp}^2\right] + \Omega_{tp}^2 + \sigma^2$
I*	$r\left[\sigma_{bt}^2 - \frac{1}{P}\Omega_{tp}^2\right] + \left[\sigma_p^2 - \frac{1}{T}\Omega_{tp}^2\right] + \Omega_{tp}^2 + \sigma^2$
R*	$\left[\sigma_p^2 - \frac{1}{T}\Omega_{tp}^2\right] + \Omega_{tp}^2 + \sigma^2$

The form of Table 22 shows just what are the estimable quantities in the analysis of variance under general conditions. We shall draw on the form of Table 22 to extend the results in the next three sections.

We can see from either Table 21 or Table 22 that the analysis of variance is, in general, not symmetric with respect to the blocks and treatments categories in a randomized block design. Only under certain rather restrictive conditions can a randomized block experiment be treated as a two-factor completely randomized experiment, i.e. when $P >> p$. Even in those circumstances it is misleading not to make a clear-cut distinction between the two for there is certainly a stronger basis for assuming that unit-treatment interactions may be neglected when the units are first classified into blocks.

Referring to either Table 21 or 22 we see that the general effect of unit-treatment interactions is to decrease, on the average, the magnitude of B*, T*, and I* relative to R*.

Evidently under the conditions of Case 3 it will not be possible in general to obtain from the analysis of variance unbiased estimates of σ_b^2, σ_t^2, σ_{bt}^2, nor of the average variance of estimates of differences of treatment effects, $t_k - t_{k'}$, when treatments are fixed ($T = t$). However, the "biasing" quantity is $\frac{1}{P} Q_{tp}^2$, and if the block size P is large then, except in pathological cases, the bias will be negligible. Then the discussion given under Case 2 applies.

We have given above some discussion of relations between components of the population model for the randomized block situation. That discussion indicates that in most cases in which we can ignore block-treatment interactions we shall also be able to treat unit-treatment interactions as negligible.

We shall therefore restrict further remarks in this section to cases in which P is not large and block-treatment interactions are not negligible.

By an unbiased error term for the estimation of, say σ_t^2, we mean some linear combination of the analysis of variance mean squares, say V_t, such that

$$E(T^* - V_t) = \sigma_t^2 \quad .$$

For the cases we consider below, unbiased error terms do not, in general, exist and we shall consider biased error terms and the bias they introduce into the estimate.

(i) $r > 1$ (i.e. R^* exists).

We might consider as an estimate for σ_t^2 the quantity

$$\hat{\sigma}_t^2 = \frac{1}{rb}\left[T^* - \frac{(B-b)}{B}(I^* - R^*) - R^*\right] \quad .$$

Then,

$$E(\hat{\sigma}_T^2) = (\sigma_t^2 - \frac{1}{BP} Q_{tp}^2)$$

and it is clear that for most combinations of B and P the bias will be negligible. Of course if $B = b$ (blocks fixed) then

$$\hat{\sigma}_t^2 = \frac{1}{rb}(T^* - R^*) \; ;$$

while if $B \gg b$ (blocks "random") then

$$\hat{\sigma}_t^2 = \frac{1}{rb}(T^* - I^*) \; .$$

To estimate σ^2_{bt} we might use

$$\hat{\sigma}^2_{bt} = \frac{1}{r}(I^* - R^*) ,$$

which has bias ($-\frac{1}{P}\Omega^2_{tp}$).

To estimate σ^2_b, we might use

$$\hat{\sigma}^2_b = \frac{1}{tr}\left[B^* - \frac{(T-t)}{T}(I^* - R^*) - R^* \right] ,$$

or

$$\tilde{\sigma}^2_b = \frac{1}{tr}\left[B^* - \frac{(T-t)}{T}(I^* - R^*) - \frac{(P-p)}{P} R^* \right] .$$

Their expectations are:

$$E(\hat{\sigma}^2_b) = \sigma^2_b + \frac{1}{P}(\sigma^2 + \sigma^2_n) ,$$

$$E(\tilde{\sigma}^2_b) = \sigma^2_b - \frac{1}{P}(\sigma^2_p) .$$

Thus, if technical errors are large compared with the unit errors one might prefer $\tilde{\sigma}^2_b$ and vice versa.

(ii) r = 1 (i.e. R* does not exist).

In this case we cannot estimate σ^2_{bt}, nor indeed $(\sigma^2 + \sigma^2_n + \sigma^2_p)$.

If B >> b, then an unbiased estimate of σ^2_t is given by $\frac{(T^* - I^*)}{b}$. Under other conditions, no linear combination of the mean squares B*, T*, I* would seem to be a reasonable estimate of σ^2_t.

If T >> t, then the estimate $\hat{\sigma}^2_b$ of σ^2_b given in (i) above becomes $\frac{1}{t}(B^* - I^*)$, and its bias is $(-\frac{1}{P}\Omega^2_{tp})$.

It is apparent that if block-treatment interactions are important then one would in general want to have treatment replication within blocks.

The conditions of Case 3 do not prevent us from obtaining unbiased estimates of linear combinations of the treatment effects when treatments are fixed (T = t). For if we define

$$\hat{t}_k = x_{.k.} - x_{...} ,$$

then, from the general statistical model appropriate for Case 3, it is easy to see that

$$E(\hat{t}_k) = t_k .$$

Hence for any set of constants $c_1, \ldots, c_T$,

$$E\left(\sum_k c_k \hat{t}_k \right) = \sum_k c_k t_k .$$

The average variance of estimated differences of treatment effects has a close connection with the analysis of variance; and under the conditions of Case 3 the average of variances of estimates such as $(\hat{t}_k - \hat{t}_{k'})$ is in fact

$$\frac{1}{rb} E(T^* - rb\sigma_t^2) = \frac{1}{rb}\left[\sigma^2 + \sigma_n^2 + \sigma_p^2 + \frac{(B-b)}{B} r\sigma_{bt}^2 - \frac{r}{P} Q_{tp}^2 \right] .$$

Thus we see that under the conditions of Case 3 the estimate of this average variance given under Case 2 will be positively biased. The effect of unit-treatment interactions is that we will tend to overestimate this average error of estimated treatment differences.

The relation of the situation on estimation of the components of variation σ_b^2, σ_t^2, and σ_{bt}^2 to that of the selection of the appropriate criteria for tests of significance is not unambiguous. It is necessary to be precise about the null hypothesis of concern.

Consider for definiteness the special case of $r > 1$, $B >> b$; i.e. blocks are random and we have treatment replication within blocks. Suppose further, for initial simplicity, that we can neglect the factor $\frac{1}{p}\Omega_{tp}^2$. We might be interested in two possible significance tests: (i) to examine the data for significance with respect to contradiction of the null assumption that treatment effects (averaged over all blocks) are equal, i.e. that $t_1 = t_2 = \ldots = t_T$; this would simply imply $\sigma_t^2 = 0$; (ii) to examine the significance of the data for contradiction of the null assumption that the treatments are all identical (while the blocks are known to differ considerably). The situation now is that if the treatments are the same then all $t_k = 0$ and also $(bt)_{ik} = 0$ for all i, j, k, so that $\sigma_t^2 = \sigma_{bt}^2 = \sigma_n^2 = \Omega_{tp}^2 = 0$.

Now for $B >> b$, $E(T^*) = E(I^*) + rb\,\sigma_t^2$. Hence a reasonable criterion for hypothesis (i) would be T^*/I^*.

However, under hypothesis (ii)

$$E(T^*) = E(I^*) = E(R^*) .$$

Any departure from the null hypothesis (ii) will occasion an inflation, on the average, in both T^* and I^*. Thus possible evaluating criteria, among others, might be

$$C_1 = \frac{T^*}{I^*} ;$$

$$C_2 = \frac{T^*}{R^*} ;$$

$$C_3 = \frac{(t-1)T^* + (t-1)(b-1)\,I^*}{b(t-1)R^*} .$$

It seems reasonable that the sensitivity of the test is related, monotonically, to the numerator and denominator degrees of freedom and to the "non-centrality factor" which we will define here as

$$m = \frac{E(\text{numerator M.S.}) - E(\text{denominator M.S.})}{E(\text{denominator M.S.})} .$$

Then putting $\sigma_0^2 = \sigma^2 + \sigma_n^2 + \sigma_p^2$, we obtain the values given in Table 23.

Table 23. Various test criteria

Criterion	m	num. d.f.	denom. d.f.
C_1	$\frac{rb\sigma_t^2}{\sigma_0^2 + r\sigma_{bt}^2}$	(t - 1)	(t - 1)(b - 1)
C_2	$\frac{r(\sigma_{bt}^2 + b\sigma_t^2)}{\sigma_0^2}$	(t - 1)	bt(r - 1)
C_3	$\frac{r(\sigma_{bt}^2 + \sigma_t^2)}{\sigma_0^2}$	b(t - 1)	bt(r - 1)

Using normal theory as a guide, one would in most cases prefer criterion C_2 for the test of equality of the treatments (hypothesis (ii)).

If in fact we cannot neglect the term $\frac{1}{P} Q_{tp}^2$, then

$$m_{C_2} = \frac{r(\sigma_{bt}^2 + b\sigma_t^2 - \frac{1}{P} Q_{tp}^2)}{\sigma_0^2}$$

$$m_{C_3} = \frac{r(\sigma_{bt}^2 + \sigma_t^2 - \frac{1}{P} Q_{tp}^2)}{\sigma_0^2} .$$

Thus the factor $\frac{1}{p} Q_{tp}^2$ does not influence the (sensitivity) outlook very much as far as the selection of a criterion for a test of the identity of treatments is concerned.

For the situation in which the test concerns $\sigma_t^2 = 0$ (i.e. equality of treatment effects averaged over all experimental units in all blocks), the considerations with respect to the test parallel those for the estimation of σ_t^2.

Whatever the null hypothesis under consideration it would seem apparent from the expected mean squares that the presence of unit-treatment interactions will occasion a decrease in sensitivity of the significance test.

9. Case 4 - factorial structure of treatments, general conditions

We consider briefly the case of a randomized block design when the treatments have a factorial structure. Results will be given for the case of two factors. The generalization to cases of more than two factors will be apparent.

Let $\mathcal{G}$ and $\mathcal{H}$ denote the factors, and let them have numbers of levels G and H, respectively, and suppose in our experiment we employ g levels of $\mathcal{G}$ and h levels of $\mathcal{H}$, randomly selected, in combination. Let m = 1, 2, . . . , G, and n = 1, 2, . . . , H denote the levels of $\mathcal{G}$ and $\mathcal{H}$, respectively, in the population of available levels. Let m* = 1, 2, . . . , g, and n* = 1, 2, . . . , h denote the randomly selected levels of $\mathcal{G}$ and of $\mathcal{H}$, respectively, in order of selection. As before, we make the convention that if G = g, then m* and m are identical indices, and similarly for n

and n* when H = h. We retain the previous notation for the blocks and units within blocks.

Connecting with the previous situation

$$GH = T;\quad gh = t;\quad p = rgh.$$

Our experiment consists of applying each of the gh selected treatment combinations to r randomly selected experimental units in each of b randomly selected blocks.

The conceptual true response from the treatment combination consisting of the m-th level of $\mathcal{G}$ and n-th level of $\mathcal{H}$ on unit j of block i we write as Y_{ijmn}. As before we take our conceptual observable as

$$y_{ijmn} = Y_{ijmn} + \epsilon_{ijmn}$$

where the ϵ_{ijmn} are taken as uncorrelated random variables, with means 0 and constant variance σ^2, reflecting technical errors.

Our population model becomes

$$y_{ijmn} = \mu + b_i + g_m + h_n + (gh)_{mn} + (bg)_{im} + (bh)_{in} + (bgh)_{imn} + p_{ij} + n_{ijmn} + \epsilon_{ijmn},$$

where $\mu = Y_{....}$

$b_i = Y_{i...} - Y_{....}$

$g_m = Y_{..m.} - Y_{....}$

$(gh)_{mn} = Y_{..mn} - Y_{..m.} - Y_{...n} + Y_{....}$

$$p_{ij} = Y_{ij..} - Y_{i...}$$

$$n_{ijmn} = Y_{ijmn} - Y_{ij..} - Y_{i.mn} + Y_{i...}$$

etc.

The statistical model for this situation can be readily developed from

$$x_{i^*m^*n^*f} = \sum_{ijmn} a_i^{i^*} \gamma_m^{m^*} \delta_n^{n^*} \rho_{i^*j}^{i^*m^*n^*f} y_{ijmn} ,$$

where $x_{i^*m^*n^*f}$ is the observation on the f-th replicate of selected treatment combination (m*n*) in selected block i*; $a_i^{i^*}$ is 1 if selected block i* corresponds to block i in the population of blocks and $a_i^{i^*}$ is 0 otherwise; $\gamma_m^{m^*}$ is 1 if selected level m* of $\mathcal{G}$ corresponds to level m in the population and $\gamma_m^{m^*}$ is 0 otherwise; etc.

The analysis of variance proceeds now according to a decomposition of the total sum of squares $G' = \sum_{i^*m^*n^*f} (x_{i^*m^*n^*f} - x_{....})^2$ into components B', G', H', I'_{GH}, I'_{BG}, I'_{BH}, I'_{BGH}, and R', where

$$B' = rgh \sum_{i^*} (x_{i^*...} - x_{....})^2$$

$$G' = rbh \sum_{m^*} (x_{..m^*.} - x_{....})^2$$

$$I'_{GH} = rb \sum_{m^*n^*} (x_{..m^*n^*} - x_{..m^*.} - x_{...n^*} + x_{....})^2$$

etc.

As before we define mean squares by replacing the ' by an *, for example

$$G^* = \frac{1}{(g-1)} G' .$$

The expectations of the analysis of variance mean squares are given in Table 24. The notational definitions are as follows:

$$\sigma_b^2 = \frac{1}{B-1} \sum_i b_i^2 ; \quad \sigma_g^2 = \frac{1}{G-1} \sum_m g_m^2 ;$$

$$\sigma_h^2 = \frac{1}{H-1} \sum_n h_n^2 ; \quad \sigma_{gh}^2 = \frac{1}{(G-1)(H-1)} \sum_{mn} (gh)_{mn}^2 ;$$

$$\sigma_{bh}^2 = \frac{1}{(B-1)(H-1)} \sum_{im} (bg)_{im}^2 ; \quad \sigma_{bg}^2 = \frac{1}{(B-1)(G-1)} \sum_{in} (bh)_{in}^2 ;$$

$$\sigma_{bgh}^2 = \frac{1}{(B-1)(G-1)(H-1)} \sum_{imn} (bgh)_{imn}^2 ; \quad \sigma_p^2 = \frac{1}{B(P-1)} \sum_{ij} p_{ij}^2 ;$$

$$\Omega_{gp}^2 = \frac{1}{B(G-1)(P-1)} \sum_{ijm} n_{ijm.}^2 ; \quad \Omega_{hp}^2 = \frac{1}{B(H-1)(P-1)} \sum_{ijn} n_{ij.n}^2 ;$$

$$\Omega_{ghp}^2 = \frac{1}{B(G-1)(H-1)(P-1)} \sum_{ijmn} (n_{ijmn} - n_{ijm.} - n_{ij.n})^2 ; \quad \sigma^2 = E(\epsilon_{ijmn}^2) ;$$

$$\Sigma_b = \sigma_b^2 - \frac{1}{G} \sigma_{bg}^2 - \frac{1}{H} \sigma_{bh}^2 - \frac{1}{P} \sigma_p^2 + \frac{1}{GH} \sigma_{bgh}^2 + \frac{1}{GP} \Omega_{gp}^2$$

$$+ \frac{1}{HP} \Omega_{hp}^2 - \frac{1}{GHP} \Omega_{ghp}^2 ;$$

$$\Sigma_g = \sigma_g^2 - \frac{1}{B} \sigma_{bg}^2 - \frac{1}{H} \sigma_{gh}^2 + \frac{1}{BH} \sigma_{bgh}^2 ;$$

$$\Sigma_h = \sigma_h^2 - \frac{1}{B} \sigma_{bh}^2 - \frac{1}{G} \sigma_{gh}^2 + \frac{1}{BG} \sigma_{bgh}^2 ;$$

$$\Sigma_{bg} = \sigma_{bg}^2 - \frac{1}{H} \sigma_{bgh}^2 - \frac{1}{P} \Omega_{gp}^2 + \frac{1}{HP} \Omega_{ghp}^2 ;$$

$$\Sigma_{bh} = \sigma_{bh}^2 - \frac{1}{G} \sigma_{bgh}^2 - \frac{1}{P} \Omega_{hp}^2 + \frac{1}{GP} \Omega_{ghp}^2 ;$$

$$\Sigma_{gh} = \sigma^2_{gh} - \frac{1}{B}\sigma^2_{bgh} ;$$

$$\Sigma_{p} = \sigma^2_{p} - \frac{1}{G}\Omega^2_{gp} - \frac{1}{H}\Omega^2_{hp} + \frac{1}{GH}\Omega^2_{ghp} ;$$

$$\Sigma_{gp} = \Omega^2_{gp} - \frac{1}{H}\Omega^2_{ghp} ;$$

$$\Sigma_{hp} = \Omega^2_{hp} - \frac{1}{G}\Omega^2_{ghp} ;$$

$$\Sigma_{ghp} = \Omega^2_{ghp} ;$$

$$\Sigma_{0} = \sigma^2 + \Sigma_{ghp} + \Sigma_{hp} + \Sigma_{gp} + \Sigma_{p} .$$

The results given in Table 24 are easily inferred from the symmetric form of the results given in Table 22. The extension to cases of more than two factors is quite immediate from Table 24.

It is apparent that all Σ quantities in Table 24 are estimable*, from the analysis of variance mean squares, when we have treatment replication within blocks (i.e. $r > 1$). In general we can find only biased estimates of components such as σ^2_g, σ^2_h, σ^2_{gh}, etc., though in many cases the bias will be negligible.

If the treatment combinations were additive with experimental units within blocks, then all Ω^2 quantities would vanish, and we could then find

*An explicit expression for the σ^2 and Ω^2 quantities in terms of the Σ's may be written down. (See for example Table 36 in Division D and some discussion in Division E, Part II.)

Table 24. Expected mean squares for two-factor experiment in randomized blocks. General conditions

Mean Square	Expected Mean Square
B^*	$rgh\Sigma_b + rh\Sigma_{bg} + rg\Sigma_{bh} + r\Sigma_{bgh} + \Sigma_0$
G^*	$rbh\Sigma_g + rh\Sigma_{bg} + rb\Sigma_{gh} + r\Sigma_{bgh} + \Sigma_0$
H^*	$rbg\Sigma_h + rg\Sigma_{bh} + rb\Sigma_{gh} + r\Sigma_{bgh} + \Sigma_0$
I^*_{BG}	$rh\Sigma_{bg} + r\Sigma_{bgh} + \Sigma_0$
I^*_{BH}	$rg\Sigma_{bh} + r\Sigma_{bgh} + \Sigma_0$
I^*_{GH}	$rb\Sigma_{gh} + r\Sigma_{bgh} + \Sigma_0$
I^*_{BGH}	$r\Sigma_{bgh} + \Sigma_0$
R^*	Σ_0

unbiased estimates for the components σ^2_g, σ^2_h, σ^2_{gh}, etc., though we would under general conditions require replication within blocks to do so.

For example, with additivity of units and treatment combinations, an unbiased estimate of σ^2_g is given by

$$\hat{\sigma}^2_g = \frac{1}{rbh}\left[G^* - \frac{(B-b)}{B} I^*_{BG} - \frac{(H-h)}{H} I^*_{GH} + \frac{(B-b)(H-h)}{BH} I^*_{BGH} - \frac{hb}{HB} R^*\right].$$

If in fact additivity of treatment combinations and units within blocks does not exist for the particular scale employed then

$$E(\hat{\sigma}^2_g) = \sigma^2_g - \frac{1}{BP} Q^2_{gp} \quad .$$

In most circumstances this bias will be negligible.

10. Case 5 - hierarchal structure of experimental material, general conditions

We consider in this section the generalization of some results when the experimental units are classified into blocks and the blocks in turn are further classified into sources, while the experimental design remains a randomized block type. For example, our treatments might be various textbooks, the sources might be various cities, the blocks within sources might be schools within the cities, the experimental units are classes within the school.

The situation we shall consider is as follows: Let there be S available sources, each containing B blocks, which in turn each contain P experimental units. Suppose we have T treatments. We select s sources at random from S; from each selected source we select b blocks at random from B; from each selected block we select rt experimental units at random from

P. We also select t treatments at random from T, and (referring to selected entities) apply treatments to experimental units at random, but in such a way that every treatment appears r times in each block.

Let the indices g, i, j, k, denote the source, block within source, unit within block within source, and treatment respectively, in the population. Thus $g = 1, 2, \ldots, S$; $i = 1, 2, \ldots, B$; $j = 1, 2, \ldots, P$; $k = 1, 2, \ldots, T$.

As before we take as our conceptual observables the random variables y_{gijk} with

$$y_{gijk} = Y_{gijk} + \epsilon_{gijk}$$

where the ϵ_{gijk} are uncorrelated random variables (reflecting possible technical errors) with means 0 and constant variance σ^2, and Y_{gijk} is the conceptual "true" response from treatment k on unit j of block i in source g.

We define

$$\mu = Y_{....}$$

$$s_g = Y_{g...} - Y_{....}$$

$$b_{gi} = Y_{gi..} - Y_{g...}$$

$$t_k = Y_{...k} - Y_{....}$$

$$(st)_{gk} = Y_{g..k} - Y_{g...} - Y_{...k} + Y_{....}$$

$$(bt)_{gik} = Y_{gi.k} - Y_{gi..} - Y_{g..k} + Y_{g...}$$

$$p_{gij} = Y_{gij.} - Y_{gi..}$$

$$n_{gijk} = Y_{gijk} - Y_{gij.} - Y_{gi.k} + Y_{gi..}$$

Each of these quantities may be given a physical interpretation. It is easily verified that the sum of all the quantities listed above is identically equal to Y_{gijk}.

Let $g^* = 1, 2, \ldots, s$; $i^* = 1, 2, \ldots, b$; $k^* = 1, 2, \ldots, t$, denote selected sources, selected blocks within selected sources, and selected treatments, respectively, all in order of their selection. However we make our usual convention that if $S = s$ then g^* and g are taken as identical indices; if $B = b$ then i^* and i are taken as identical indices; if $T = t$ then k^* and k are taken as identical indices.

Let $x_{g^*i^*k^*f}$ denote the f-th replicate of selected treatment k^* in selected block i^* of selected source g^* where $f = 1, 2, \ldots, r$ for each $(g^*i^*k^*)$. To write an explicit model for $x_{g^*i^*k^*f}$ we define the following random variables:

Let $a_g^{g^*}$ = 1 if selected source g^* corresponds to source g in the population,

= 0 otherwise;

$\beta_{g^*i}^{g^*i^*}$ = 1 if selected block i^* in selected source g^* corresponds to block i in the population of blocks in selected source g^*,

= 0 otherwise;

$\gamma_k^{k^*}$ = 1 if selected treatment k^* corresponds to treatment k in the population,

= 0 otherwise;

$\rho_{g^*i^*j}^{g^*i^*k^*f}$ = 1 if the f-th replicate of selected treatment k^* in selected block i^* of selected source g^* falls on unit j of selected block i^* of selected source g^*,

= 0 otherwise.

The statistical model may be developed from

$$x_{g*i*k*f} = \sum_{gijk} a_g^{g*} \beta_{g*i}^{g*i*} \gamma_k^{k*} \rho_{g*i*j}^{g*i*k*f} (Y_{gijk} + \epsilon_{gijk}) ,$$

using the population parameters μ, s_g, b_{gi}, t_k, etc. defined above.

The analysis of variance involves the decomposition of the total sum of squares

$$\sum_{g*i*k*f} (x_{g*i*k*f} - x_{....})^2$$

into components S', B', T', I'_{ST}, I'_{BT}, and R', with degrees of freedom (s-1), s(b-1), (t-1), (s-1)(t-1), s(b-1)(t-1), and sbt(r-1) respectively, where

$$S' = btr \sum_{g*} (x_{g*...} - x_{....})^2 ,$$

$$B' = tr \sum_{g*i*} (x_{g*i*..} - x_{g*...})^2 ,$$

$$T' = sbr \sum_{k*} (x_{..k*.} - x_{....})^2 ,$$

etc.

As before we denote mean squares by replacing a ' by an *, for example,

$$S* = \frac{1}{(s-1)} S' .$$

The expectations of the analysis of variance mean squares are given in Table 25. The results of Table 25 are easily inferred from the symmetric form of results for the simpler case given in Table 22. The notational definitions used in Table 25 are as follows:

$$\sigma_s^2 = \frac{1}{S-1} \sum_g s_g^2 ; \quad \sigma_b^2 = \frac{1}{S(B-1)} \sum_{gi} b_{gi}^2 ;$$

$$\sigma_t^2 = \frac{1}{T-1} \sum_k t_k^2 ; \quad \sigma_{st}^2 = \frac{1}{(S-1)(T-1)} \sum_{gk} (st)_{gk}^2 ;$$

$$\sigma^2_{bt} = \frac{1}{S(B-1)(T-1)} \sum_{gik} (bt)^2_{gik}; \quad \sigma^2_p = \frac{1}{SB(P-1)} \sum_{gij} e^2_{gij} ;$$

$$Q^2_{tp} = \frac{1}{SB(P-1)(T-1)} \sum_{gijk} n^2_{gijk}; \quad \sigma^2 = E(\epsilon^2_{gijk}) ;$$

$$\Sigma_s = \sigma^2_s - \frac{1}{T} \sigma^2_{st} - \frac{1}{B} \sigma^2_b + \frac{1}{TB} \sigma^2_{bt} ;$$

$$\Sigma_b = \sigma^2_b - \frac{1}{T} \sigma^2_{bt} - \frac{1}{P} \sigma^2_p + \frac{1}{TP} Q^2_{tp} ;$$

$$\Sigma_p = \sigma^2_p - \frac{1}{T} Q^2_{tp} ;$$

$$\Sigma_t = \sigma^2_t - \frac{1}{S} \sigma^2_{ts} ;$$

$$\Sigma_{st} = \sigma^2_{st} - \frac{1}{B} \sigma^2_{bt} ;$$

$$\Sigma_{bt} = \sigma^2_{bt} - \frac{1}{P} Q^2_{tp} ;$$

$$\Sigma_{tp} = Q^2_{tp} ;$$

$$\Sigma_0 = \sigma^2 + \Sigma_{tp} + \Sigma_p .$$

The extension of Table 25 to cover cases of multiple nesting of experimental material should be readily apparent. The results of Tables 24 and 25 may be combined to indicate the expected mean squares for multiple factorial structures of treatments together with multiple nesting

Table 25. Expected mean squares for hierarchal structure of experimental material in randomized block design. General conditions

Mean Square	Expected Mean Square
S^*	$btr\Sigma_s + br\Sigma_{st} + tr\Sigma_b + r\Sigma_{bt} + \Sigma_0$
B^*	$tr\Sigma_b + r\Sigma_{bt} + \Sigma_0$
T^*	$sbr\Sigma_t + br\Sigma_{st} + r\Sigma_{bt} + \Sigma_0$
I^*_{ST}	$br\Sigma_{st} + r\Sigma_{bt} + \Sigma_0$
I^*_{BT}	$r\Sigma_{bt} + \Sigma_0$
R^*	Σ_0

of experimental material. The similarity of the pattern of results in Tables 24 and 25 to those for the simpler situation in Table 22 will be evident on inspection. A main objective of the last two sections was to bring to the fore the basic patterns of Table 22.

11. Case 6 - hierarchal structure of treatments, general conditions

The situation we consider in this section is one which is conceptually somewhat different from those considered heretofore, in that our treatments will have a hierarchal structure. The design framework we employ is the randomized block.

Suppose we have CT treatments categorized into C categories each containing T treatments. (An example might be seeds classified as varieties.) Let m = 1, 2, . . . , C and n = 1, 2, . . . , T denote category ($\mathcal{C}$) and treatment within category ($\mathcal{T}$), respectively.

Suppose as before we have B blocks each of P experimental units. Let i = 1, 2, . . . , B, and j = 1, 2, . . . , P denote block ($\mathcal{B}$) and unit within block, respectively.

We select c categories from C at random, and then select t = rb treatments at random from T within each selected category. We select at random b blocks from B, and ct = crb units from P in each selected block. We then apply treatments to units at random but with the restriction that every selected category is represented on r units in every selected block. Let i* = 1, 2, . . . , b denote the selected blocks, in order of selection. Let m* = 1, 2, . . . , c denote selected categories, in order of selection. We make our usual convention that if B = b, then i* and i are the same index; if C = c then m* and m are the same index.

If we had no basis or desire for identifying individual treatments within a category then we might regard the category as designating an "ideal treatment", and treat the variability of individuals within categories as technical error. This is the point of view favored elsewhere. With this latter point of view the present Case 6 becomes identical with the situation studied in Cases 1, 2, and 3.

We shall, however, take the treatments within categories as individuals in this section and from this point of view we admit at the outset

the impossibility of replication of individual treatments.

Let Y_{ijmn} denote the conceptual "true" response for a given scale if treatment n of category m were applied to unit j of block i. For our conceptual observable we take

$$y_{ijmn} = Y_{ijmn} + \epsilon_{ijmn},$$

where the ϵ_{ijmn} represent errors of measurement, etc., and are assumed to be uncorrelated random variables having means 0 and constant variance σ^2 (though this last assumption could be relaxed somewhat with very little complication).

Then

$\mu = Y_{....}$ is the overall true response from all treatments of all categories on all experimental units;

$b_i = Y_{i...} - \mu$ is the effect of block i;

$p_{ij} = Y_{ij..} - Y_{i...}$ is the within i-th block effect of unit j of block i;

$c_m = Y_{..m.} - \mu$ is the effect of category m;

$t_{mn} = Y_{..mn} - Y_{..m.}$ is the within m-th category effect of treatment n of category m;

$(bc)_{im} = Y_{i.m.} - Y_{i...} - Y_{..m.} + \mu$ is the interaction of the i-th block and m-th category;

$(bt)_{imn} = (Y_{i.mn} - Y_{i.m.} - Y_{..mn} + Y_{..m.})$ is the within m-th category interaction of treatment n of category m with block i.

$(pc)_{ijm} = (Y_{ijm.} - Y_{i.m.} - Y_{ij..} + Y_{i...})$ is the within i-th block interaction of unit j of block i with category m.

$(pt)_{ijmn} = (Y_{ijmn} - Y_{i.mn} - Y_{ijm.} + Y_{i.m.})$ is the within i-th block and m-th category interaction of unit j of block i with treatment n of category m.

It is easily checked that the sum of all the quantities defined above is identically Y_{ijmn}.

Let $x_{i^*m^*f}$ denote the observation from the f-th repetition of selected category m* in selected block i*, where f = 1, 2, . . . , r for each (i*m*). To write a model for the observations we define the following random variables:

Let $\alpha_i^{i^*}$ = 1 if selected block i* corresponds to block i in the population of blocks,

= 0 otherwise;

$\beta_m^{m^*}$ = 1 if selected category m* corresponds to category m in the population of categories,

= 0 otherwise;

$\delta_{i^*j}^{i^*m^*f}$ = 1 if the f-th repetition of selected category m* in block i* falls on unit j of selected block i*,

= 0 otherwise;

$\rho_{m^*n}^{i^*m^*f}$ = 1 if the f-th repetition of selected category m* in block i* corresponds to the n-th treatment in selected category m*,

= 0 otherwise.

Some of the distributional properties of these random variables are as follows:

The α's, β's, δ's and ρ's are groupwise statistically independent.

$$E(\alpha_i^{i^*}) = \frac{1}{B} .$$

$$E(\alpha_i^{i^*} \alpha_{i'}^{i^{*\prime}}) = \frac{1}{B(B-1)}, \quad i \neq i', \; i^* \neq i^{*\prime} .$$

$$E(\beta_m^{m*}) = \frac{1}{C}$$

$$E(\beta_m^{m*}\beta_{m'}^{m*'}) = \frac{1}{C(C-1)}$$

$$E(\delta_{i*j}^{i*m*f}) = \frac{1}{P}$$

$$E(\delta_{i*j}^{i*m*f}\ \delta_{i*j'}^{i*m*'f'}) = \frac{1}{P(P-1)},\ (m*f) \neq (m*'f'),\ j \neq j' .$$

For $i* \neq i*'$ the δ's are independent.

$$E(\rho_{m*n}^{i*m*f}) = \frac{1}{T}$$

$$E(\rho_{m*n}^{i*m*f}\ \rho_{m*n'}^{i*m*'f'}) = \frac{1}{T(T-1)},\ (i*f) \neq (i*'f'),\ n \neq n' .$$

For $m* \neq m*'$, the ρ's are independent.

The statistical model for the observations is then

$$x_{i*m*f} = \mu + \sum_i \alpha_i^{i*} b_i + \sum_m \beta_m^{m*} c_m + \sum_{im} \alpha_i^{i*}\beta_m^{m*}(bc)_{im} + \sum_{ij}\alpha_i^{i*}\delta_{i*j}^{i*m*f} p_{ij}$$

$$+ \sum_{mn} \beta_m^{m*}\rho_{m*n}^{i*m*f} t_{mn} + \sum_{ijm} \alpha_i^{i*}\beta_m^{m*}\delta_{i*j}^{i*m*f}(pc)_{ijm}$$

$$+ \sum_{imn} \alpha_i^{i*}\beta_m^{m*}\rho_{m*n}^{i*m*f}(bt)_{imn} + \sum_{ijmn} \alpha_i^{i*}\beta_m^{m*}\delta_{i*j}^{i*m*f}$$

$$((pt)_{ijmn} + \epsilon_{ijmn}) .$$

We note that the experimental design was such that each selected category was represented r times in each selected block, but the selected individuals within a selected category were randomized over blocks as well as over units within blocks.

The usual randomized block analysis of variance would partition

$$\sum_{i*m*f} (x_{i*m*f} - x_{...})^2$$

into B', C', I'_{BC}, R', with (b-1), (c-1), (b-1)(c-1), and bc(r-1) degrees of freedom respectively. Denoting the mean squares by replacing the ' by an *, we would get the expected mean squares given in Table 26. The notational definitions are as follows:

$$\Sigma_b = \sigma_b^2 - \frac{1}{C}\sigma_{bc}^2 - \frac{1}{P}\sigma_p^2 + \frac{1}{CP}\sigma_{pc}^2 ,$$

$$\Sigma_c = \sigma_c^2 - \frac{1}{B}\sigma_{bc}^2 - \frac{1}{T}\sigma_t^2 + \frac{1}{BT}\sigma_{bt}^2 ,$$

$$\Sigma_p = \sigma_p^2 - \frac{1}{C}\sigma_{pc}^2 ,$$

$$\Sigma_t = \sigma_t^2 - \frac{1}{B}\sigma_{bt}^2 ,$$

$$\Sigma_{bc} = \sigma_{bc}^2 - \frac{1}{P}\sigma_{pc}^2 - \frac{1}{T}\sigma_{bt}^2 + \frac{1}{PT}\sigma_{pt}^2 ,$$

$$\Sigma_{bt} = \sigma_{bt}^2 - \frac{1}{P}\sigma_{pt}^2 ,$$

$$\Sigma_{pc} = \sigma_{pc}^2 - \frac{1}{T}\sigma_{pt}^2 ,$$

$$\Sigma_{pt} = \sigma_{pt}^2,$$

$$\Sigma_0 = \sigma^2 + \Sigma_{pt} + \Sigma_{pc} + \Sigma_{bt} + \Sigma_p + \Sigma_t,$$

and the σ^2 quantities are defined as usual, for example,

$$\sigma_b^2 = \frac{1}{B-1}\sum_i b_i^2, \quad \sigma_{bt}^2 = \frac{1}{(B-1)C(T-1)}\sum_{imn} (bt)_{imn}^2, \quad \text{etc.}$$

Table 26. Expected mean squares for randomized block design with hierarchal structure of treatments. Case 6

Mean Square	Expected Mean Square
B*	$rc\Sigma_b + r\Sigma_{bc} + \Sigma_0$
C*	$rb\Sigma_c + r\Sigma_{bc} + \Sigma_0$
I^*_{BC}	$r\Sigma_{bc} + \Sigma_0$
R*	$\Sigma_0 = \sigma^2 + \Sigma_{pt} + \Sigma_{pc} + \Sigma_{bt} + \Sigma_p + \Sigma_t$

Direct use may be made of Table 26 in the form given by means of the following relationships:

$$\sigma_b^2 = \Sigma_b + \frac{1}{C}\Sigma_{bc} + \frac{1}{P}\Sigma_p + \frac{1}{CT}\Sigma_{bt} + \frac{1}{CP}\Sigma_{pc} + \frac{1}{CTP}\Sigma_{pt}$$

$$\sigma_c^2 = \Sigma_c + \frac{1}{B}\Sigma_{bc} + \frac{1}{T}\Sigma_t + \frac{1}{BP}\Sigma_{pc} + \frac{1}{BT}\Sigma_{bt} + \frac{1}{BPT}\Sigma_{pt}$$

$$\sigma_{bc}^2 = \Sigma_{bc} + \frac{1}{P}\Sigma_{pc} + \frac{1}{T}\Sigma_{bt} + \frac{1}{PT}\Sigma_{pt}$$

$$\sigma_t^2 = \Sigma_t + \frac{1}{B}\Sigma_{bt} + \frac{1}{BP}\Sigma_{pt}$$

$$\sigma_p^2 = \Sigma_p + \frac{1}{C}\Sigma_{pc} + \frac{1}{CT}\Sigma_{pt}$$

$$\sigma_{bt}^2 = \Sigma_{bt} + \frac{1}{P}\Sigma_{pt}$$

$$\sigma_{pc}^2 = \Sigma_{pc} + \frac{1}{T}\Sigma_{pt}$$

$$\sigma_{pt}^2 = \Sigma_{pt}.$$

The residual mean square R* might be regarded as a mean square for treatments within blocks and categories. There is however no error term for the evaluation of effects of treatments within categories.

As an estimate of, for example, σ_c^2 we might consider

$$\hat{\sigma}_c^2 = \hat{\Sigma}_c + \frac{1}{B}\hat{\Sigma}_{bc}$$

where $\hat{\Sigma}_c = \frac{1}{rb}(C^* - I^*_{BC})$,

$$\hat{\Sigma}_{bc} = \frac{1}{r}(I^*_{BC} - R^*) .$$

Then

$$E(\hat{\sigma}_c^2) = \Sigma_c + \frac{1}{B}\Sigma_{bc}$$

$$= \sigma_c^2 - \frac{1}{T}\sigma_t^2 - \frac{1}{BP}\sigma_{pc}^2 + \frac{1}{BPT}\sigma_{pt}^2 .$$

C. The Latin Square Design

The analysis of variance of an experiment in a latin square design is usually justified by way of general linear hypothesis theory, using an assumed linear model

observation = (row effect) + (column effect) + (treatment effect) + (random error)

where the random errors are taken as normally and independently distributed with means 0 and constant variance. From this model, estimates, reliability of estimates and tests of significance or tests of hypothesis can be derived.

Fisher (1926) and Fisher and Yates (1935, 1938) have laid down the rules which experimenters should follow in setting up latin square experiments, including rather complicated (at least from some points of view) rules about the randomization to be followed.

The question arises of the relationship of randomization procedures to the assumptions used and also of the physical meaning to be attached to the components of the model. Some of these matters have been investigated by Kempthorne (1952a), who has discussed the analysis of the latin square design under the assumption of additivity of treatments and experimental units. This assumption is that the response of treatment k on the unit in the i-th row and j-th column is given by $(t_k + x_{ij})$, where t_k is independent of the unit and x_{ij} is independent of the treatment.

One objective of the present work is to examine the latin square design under randomization without the restrictive additivity assumptions. At the same time the situation studied is stated in general terms so as to include so-called "fixed models" and "random models" as special cases.

General discussions of the latin square and its use in statistical design of experiments have been given by Fisher (1926, 1935a), Fisher and Yates (1938), Yates (1936), Cochran and Cox (1950) and in most text books on statistical methods and experimental design.

1. The experimental situation and design

We suppose that we have available experimental material divided into units, each of which is classified according to two criteria, which we shall refer to as R (row) and C (column), in such a way that specifying

row and column specifies a unit uniquely. For examples: a field may be divided into plots according to a rectangular grid system, and then each plot classified according to the row and column of the grid; a sheet of paper coming off a roller in a papermaking factory can be considered to be divided into rectangular units, each unit classified as to distance from lateral edge and also as to distance along the sheet produced. Let R be the number of levels of $\mathcal{R}$ and C be the number of levels of $\mathcal{C}$, giving RC units in all.

Suppose now that we have T treatments which we wish to study with respect to this experimental material. These treatments may be a full or partial factorial set of any number of factors such as machines, operators, temperatures, compositions, catalysts, etc., or may be non-factorial in relationship.

Our experimental procedure is to select t rows at random from R, t columns at random from C, thus giving us a "square" of t^2 units. We then select t treatments at random from T, and apply treatments to experimental units at random in a latin square arrangement according to the rules given by Fisher and Yates (1938). In the sequel we shall make explicit use of the property, stated by Fisher (1926) as a requirement of the latin square design, that every pair of plots not in the same row or column belong to the same treatment equally frequently.

2. The population and population model

Let $i = 1, 2, \ldots, R$ denote the row and $j = 1, 2, \ldots, C$ denote the column, in the original classification of the experimental material. Then every unit is identified by a number pair (i, j).

Let $k = 1, 2, \ldots, T$ denote the treatment in the set of treatments of interest.

We assume that the response from any treatment k on unit (i, j) would depend only on that unit and that treatment, and would be independent of which treatments were applied to other experimental units. In a field experiment the observed yields of a plot may depend not only on the plot and the treatment applied thereto but also on the actual configuration of treatments as applied in the whole experiment. Such a case will not be discussed at all herein. However, we know that if we attempted to apply treatment k to unit (i, j) there would be variations in treatment application and errors in measurement of response. We shall suppose that our conceptual observable quantity, if we applied treatment k to unit (i, j), is a random variable y_{ijk} such that

$$y_{ijk} = Y_{ijk} + \epsilon_{ijk},$$

where Y_{ijk} is the "true" response from treatment k on unit (i, j) and ϵ_{ijk} is a random variable with 0 mean, variance σ^2. Further we shall assume that the ϵ_{ijk} are uncorrelated.

We now define some parameters of the population of "true" responses $\{Y_{ijk}\}$, a population of RCT numbers, and give these parameters a physical interpretation. We use the usual dot convention to denote means.

$\mu = Y_{...}$ is the overall mean "true" response if every treatment were applied to every experimental unit.

$r_i = Y_{i..} - Y_{...}$ is the difference between the average of responses of all treatments on all units of row i and μ. We refer to r_i as the effect of row i.

Similarly $c_j = Y_{.j.} - Y_{...}$ is the effect of column j.

$t_k = Y_{..k} - Y_{...}$ is the difference between the mean response from treatment k over all the experimental units and μ. We call t_k the effect of the k-th treatment.

$(rt)_{ik} = Y_{i.k} - Y_{i..} - Y_{..k} + Y_{...}$ measures the difference between the effect of treatment k within the i-th row and its effect over all the rows. We shall refer to $(rt)_{ik}$ as the interaction of row i and treatment k.

Similarly $(ct)_{jk} = Y_{.jk} - Y_{.j.} - Y_{..k} + Y_{...}$ is the interaction of column j and treatment k.

Similarly $e_{ij} = Y_{ij.} - Y_{i..} - Y_{.j.} + Y_{...}$ is the interaction of row i and column j. The set $\{e_{ij}\}$ act as additive errors with respect to comparisons of treatments in the analysis of the eventual design. This is the reason why we have used the notation e_{ij} rather than $(rc)_{ij}$.

$n_{ijk} = Y_{ijk} - Y_{i.k} - Y_{.jk} + Y_{..k} - Y_{ij.} + Y_{i..} + Y_{.j.} - Y_{...}$ represents a three-way interaction of row i, column j and treatment k. A more consistent, but also more troublesome, notation would be $(rct)_{ijk}$.

Hence the model for the conceptual observables is

$$y_{ijk} = \mu + r_i + c_j + t_k + (rt)_{ik} + (ct)_{jk} + e_{ij} + n_{ijk} + \epsilon_{ijk} .$$

This expression we call the population model. The components of this model have a defined relationship to the physical situation.

3. Relationships among components of the population model

It is of interest to examine some of the intrinsic properties of components of the population model, as well as their relationships in the light of the experimental situation.

First we note that, by definition,

$$\sum_i r_i = \sum_j c_j = \sum_k t_k = \sum_i (rt)_{ik} = \sum_k (rt)_{ik} = \sum_j (ct)_{jk} = \sum_k (ct)_{jk} = \sum_i e_{ij}$$

$$= \sum_j e_{ij} = \sum_i n_{ijk} = \sum_j n_{ijk} = \sum_k n_{ijk} = 0.$$

If the treatments are all identical then $Y_{ij.} = Y_{ijk}$ for all k, and then $t_k = (rt)_{ik} = (ct)_{jk} = n_{ijk} = 0$, for all i, j, k. (This is the null hypothesis considered by Fisher (1935b)).

If all the units of each row are identical, then $Y_{ijk} = Y_{i.k}$, for all i, j and k, and then $c_j = (ct)_{jk} = e_{ij} = n_{ijk} = 0$.

If all the units in each column are identical, then $Y_{ijk} = Y_{.jk}$ for all i, j, and k, and then $r_i = (rt)_{ik} = e_{ij} = n_{ijk} = 0$.

The common feature of each of these special cases is that the $n_{ijk} = 0$. Moreover the n_{ijk} represent interactions of three classifications. We might expect that in many cases the n_{ijk} would tend to be small.

For a more formal view of the relationships we might consider the true response Y_{ijk} to be a function of underlying variables describing properties of the experimental units and of the treatments, say

$$Y_{ijk} = f(w_{ij}, z_k),$$

where w_{ij} describes a property of unit (i, j) and z_k describes a property of treatment k. Writing a Taylor's expansion of Y_{ijk}, about the point $(w_{..}, z_.)$, we have, up to terms of order 1/2* ,

$$Y_{ijk} \doteq \bar{f} + u_{ij}\bar{f}_1 + v_k\bar{f}_2 + u_{ij}^2 \frac{\bar{f}_{11}}{2} + v_k^2 \frac{\bar{f}_{22}}{2} + u_{ij}v_k\bar{f}_{12}$$

$$+ u_{ij}^2 v_k \frac{\bar{f}_{112}}{2} + u_{ij} v_k^2 \frac{\bar{f}_{122}}{2} ,$$

where $\bar{f} = f(w_{..}, z_.)$;

$$\bar{f}_1 = \frac{\partial f}{\partial w}\Bigg|_{\substack{w = w_{..} \\ z = z_.}} ; \quad \bar{f}_2 = \frac{\partial f}{\partial z}\Bigg|_{\substack{w = w_{..} \\ z = z_.}} ; \quad \bar{f}_{12} = \frac{\partial^2 f}{\partial z \partial w}\Bigg|_{\substack{w = w_{..} \\ z = z_.}} ; \text{ etc.}$$

$$u_{ij} = w_{ij} - w_{..} ;$$

$$v_k = z_k - z_. .$$

Then, putting $U_{i.}^2 = \frac{1}{C} \sum_j u_{ij}^2$, $U_{.j}^2 = \frac{1}{R} \sum_i u_{ij}^2$; $U_{..}^2 = \frac{1}{RC} \sum_{ij} u_{ij}^2$, and $V^2 = \frac{1}{T} \sum_k v_k^2$, we have, approximately,

$$Y_{ij.} \doteq \bar{f} + u_{ij}\bar{f}_1 + u_{ij}^2 \frac{\bar{f}_{11}}{2} + V^2 \frac{\bar{f}_{22}}{2} + u_{ij} V^2 \frac{\bar{f}_{122}}{2}$$

$$Y_{i.k} \doteq \bar{f} + u_{i.}\bar{f}_1 + v_k\bar{f}_2 + U_{i.}^2 \frac{\bar{f}_{11}}{2} + v_k^2 \frac{\bar{f}_{22}}{2} + u_{i.}v_k\bar{f}_{12}$$

$$+ U_{i.}^2 v_k \frac{\bar{f}_{112}}{2} + u_{i.} v_k^2 \frac{\bar{f}_{122}}{2}$$

$$Y_{.jk} \doteq \bar{f} + u_{.j}\bar{f}_1 + v_k\bar{f}_2 + U_{.j}^2 \frac{\bar{f}_{11}}{2} + v_k^2 \frac{\bar{f}_{22}}{2} + u_{.j}v_k\bar{f}_{12}$$

$$+ U_{.j}^2 v_k \frac{\bar{f}_{112}}{2} + u_{.j} v_k^2 \frac{\bar{f}_{122}}{2}$$

---

*It is common to make a Taylor expansion up to a particular degree of partial derivatives. We have not followed that procedure since the importance of a term may depend more on the divisor than on the degree of the derivative. This point has been suggested independently by Tick (1954).

$$Y_{i..} \doteq \bar{f} + u_{i.}\bar{f}_1 + U^2_{i.}\frac{\bar{f}_{11}}{2} + V^2\frac{\bar{f}_{22}}{2} + u_{i.}V^2\frac{\bar{f}_{122}}{2}$$

$$Y_{.j.} \doteq \bar{f} + u_{.j}\bar{f}_1 + U^2_{.j}\frac{\bar{f}_{11}}{2} + V^2\frac{\bar{f}_{22}}{2} + u_{.j}V^2\frac{\bar{f}_{122}}{2}$$

$$Y_{..k} \doteq \bar{f} + v_k\bar{f}_2 + U^2_{..}\frac{\bar{f}_{11}}{2} + v_k^2\frac{\bar{f}_{22}}{2} + U^2_{..}v_k\frac{\bar{f}_{112}}{2}$$

$$Y_{...} \doteq \bar{f} + U^2_{..}\frac{\bar{f}_{11}}{2} + V^2\frac{\bar{f}_{22}}{2}$$

Thus

$$r_i \doteq u_{i.}\bar{f}_1 + (U^2_{i.} - U^2_{..})\frac{\bar{f}_{11}}{2} + u_{i.}V^2\frac{\bar{f}_{22}}{2}$$

$$c_j \doteq u_{.j}\bar{f}_1 + (U^2_{.j} - U^2_{..})\frac{\bar{f}_{11}}{2} + u_{.j}V^2\frac{\bar{f}_{122}}{2}$$

$$t_k \doteq v_k\bar{f}_2 + (v_k^2 - V^2)\frac{\bar{f}_{22}}{2} + U^2_{..}v_k\frac{f_{112}}{2}$$

$$e_{ij} \doteq (u_{ij}-u_{i.}-u_{.j})\bar{f}_1 + (u^2_{ij}-U^2_{i.}-U^2_{.j}+U^2_{..})\frac{\bar{f}_{11}}{2} + (u_{ij}-u_{i.}-u_{.j})V^2\frac{\bar{f}_{122}}{2}$$

$$(rt)_{ik} \doteq u_{i.}v_k\bar{f}_{12} + (U^2_{i.}-U^2_{..})v_k\frac{\bar{f}_{112}}{2} + u_{i.}(v_k^2-V^2)\frac{\bar{f}_{122}}{2}$$

$$(ct)_{jk} \doteq u_{.j}v_k\bar{f}_{12} + (U_{.j}-U^2_{..})v_k\frac{\bar{f}_{112}}{2} + u_{.j}(v_k^2 - V^2)\frac{\bar{f}_{122}}{2}$$

$$n_{ijk} \doteq (u_{ij}-u_{i.}-u_{.j})v_k\bar{f}_{12} + (u^2_{ij}-U^2_{i.}-U^2_{.j}+U^2_{..})v_k\frac{\bar{f}_{112}}{2}$$
$$+ (u_{ij}-u_{i.}-u_{.j})(v_k^2 - V^2)\frac{\bar{f}_{122}}{2} \quad .$$

To the degree of approximation of the expansion given, we see that n_{ijk} depends on the magnitude of higher order derivatives than does e_{ij}, and further that if the e_{ij} are negligible then the coefficients $(u_{ij}-u_{i.}-u_{.j})$ and $(u^2_{ij}-U^2_{i.}-U^2_{.j}+U^2_{..})$ must be negligible and hence the n_{ijk} must be

negligible. However, if the n_{ijk} are negligible then this may be due to $\bar{f}_{12}$, $\bar{f}_{112}$ and $\bar{f}_{122}$ being small. In general the magnitude of the n_{ijk} will be less than that of $(rt)_{ik}$ and $(ct)_{jk}$ for $(u_{ij}-u_{i.}-u_{.j})$ will generally be smaller in absolute value than u_{ij}, $u_{i.}$ or $u_{.j}$; similarly for $(u^2_{ij}-U^2_{i.}-U^2_{.j}+U^2_{..})$ in relation to $(U^2_{i.}-U^2_{..})$ and $(U^2_{.j}-U^2_{..})$.

Another type of approximation which one might consider is that the RC variables $\{w_{ij}\}$ can be adequately described by (R+C) variables via a function $w_{ij} = h(a_i, b_j)$, where a_i and b_j are row and column properties respectively. Then we would have

$$Y_{ijk} \doteq g(a_i, b_j, z_k).$$

Writing a Taylor's expansion about $(a_.\,, b_.\,, z_.)$, to the same order as before, we obtain

$$\begin{aligned} Y_{ijk} \doteq \bar{g} &+ \bar{g}_1 A_i + \bar{g}_2 B_j + \bar{g}_3 Z_k + \frac{\bar{g}_{11}}{2} A_i^2 + \frac{\bar{g}_{22}}{2} B_j^2 + \frac{\bar{g}_{33}}{2} Z_k^2 \\ &+ \bar{g}_{12} A_i B_j + \bar{g}_{13} A_i Z_k + \bar{g}_{23} B_j Z_k + \bar{g}_{123} A_i B_j Z_k \\ &+ \frac{\bar{g}_{112}}{2} A_i^2 B_j + \frac{\bar{g}_{113}}{2} A_i^2 Z_k + \frac{\bar{g}_{122}}{2} A_i B_j^2 + \frac{\bar{g}_{223}}{2} B_j^2 Z_k \\ &+ \frac{\bar{g}_{133}}{2} A_i Z_k^2 + \frac{\bar{g}_{233}}{2} B_j Z_k^2 \ , \end{aligned}$$

where $\bar{g} = g(a_.\,, b_.\,, z_.)$;

$$\bar{g}_1 = \frac{\partial g}{\partial a} \Bigg|_{\substack{a = a_. \\ b = b_. \\ z = z_.}}$$

$$\bar{g}_{12} = \frac{\partial^2 g}{\partial a \partial b} \Bigg|_{\substack{a = a_. \\ b = b_. \\ z = z_.}} ; \qquad \text{etc.};$$

$A_i = (a_i - a_.)$; $B_j = (b_j - b_.)$; $Z_k = (z_k - z_.)$.

Putting $A^2 = \frac{1}{R} \sum_i A_i^2$; $B^2 = \frac{1}{C} \sum_j B_j^2$; $Z^2 = \frac{1}{T} \sum_k Z_k^2$, we find that

$$r_i \doteq A_i \bar{g}_1 + (A_i^2 - A^2) \frac{\bar{g}_{11}}{2} + A_i B^2 \frac{\bar{g}_{122}}{2} + A_i Z^2 \frac{\bar{g}_{133}}{2} ;$$

$$c_j \doteq B_j \bar{g}_2 + (B_j^2 - B^2) \frac{\bar{g}_{22}}{2} + A^2 B_j \frac{\bar{g}_{112}}{2} + B_j Z^2 \frac{\bar{g}_{233}}{2} ;$$

$$t_k \doteq Z_k \bar{g}_3 + (Z_k^2 - Z^2) \frac{\bar{g}_{33}}{2} + A^2 Z_k \frac{\bar{g}_{113}}{2} + B^2 Z_k \frac{\bar{g}_{223}}{2} ;$$

$$e_{ij} \doteq A_i B_j \bar{g}_{12} + (A_i^2 - A^2) B_j \frac{\bar{g}_{112}}{2} + A_i (B_j^2 - B^2) \frac{\bar{g}_{122}}{2} ;$$

$$n_{ijk} \doteq A_i B_j Z_k \bar{g}_{123} ;$$

etc.

Under these conditions it is again obvious that if the e_{ij} are small so also will be the n_{ijk}, in general, but conversely the n_{ijk} may be negligible while the e_{ij} are large.

As an example of the application of these ideas consider the situation

$$Y_{ijk} = K + a_i b_j + (a_i + b_j) z_k .$$

Then $\bar{g}_1 = b_. + z_.$

$\bar{g}_2 = a_. + z_.$

$\bar{g}_{12} = 1$

$\bar{g}_{13} = 1$

$$\bar{g}_{12} = 1$$

$$\bar{g}_{123} = 0.$$

Thus, in this case, $n_{ijk} = 0$, while the e_{ij}, $(rt)_{ik}$, $(ct)_{jk}$ would not be 0.

However, if the relationship were, say,

$$Y_{ijk} = K + z_k^{(a_i+b_j)}$$

then
$$\bar{g}_1 = (\log z_{\cdot}) z_{\cdot}^{(a_{\cdot}+b_{\cdot})}$$

$$\bar{g}_{12} = (\log z_{\cdot})^2 z_{\cdot}^{(a_{\cdot}+b_{\cdot})}$$

$$\bar{g}_{123} = (\log z_{\cdot})^2 z_{\cdot}^{(a_{\cdot}+b_{\cdot}-1)} + \frac{2}{z_{\cdot}} (\log z_{\cdot}) z_{\cdot}^{(a_{\cdot}+b_{\cdot})}$$

$$= (\log z_{\cdot})(\log z_{\cdot} + 2) z_{\cdot}^{(a_{\cdot}+b_{\cdot}-1)},$$

whence we cannot conclude that the n_{ijk} would necessarily be small.

4. The statistical model

We have defined our population as a set of RCT (conceptual) true responses $\{Y_{ijk}\}$, have defined certain parameters of this population and have given these a physical interpretation in terms of the experimental situation. To take into consideration technical errors in applying treatments and making measurements we have assumed our population of conceptual observables to be a set $\{y_{ijk}\}$, where $y_{ijk} = Y_{ijk} + \epsilon_{ijk}$, and the ϵ_{ijk} are mutually uncorrelated, have means 0 and constant variance σ^2. This assumption of additivity of true response and of technical errors is not trivial, but for a general study we have no reasonable recourse but assumption for this aspect. (The homogeneity

assumption could of course be relaxed somewhat with little complication.) The assumption may be reasonable in many experimental circumstances.

The experiment we carry out according to the procedure described in a preceding section may now be regarded as giving us a random (within the restrictions of the design) sample from the population of conceptual observables $\{y_{ijk}\}$. We proceed now to develop notation for this sample.

Let $i^* = 1, 2, \ldots, t$; $j^* = 1, 2, \ldots, t$; $k^* = 1, 2, \ldots, t$ denote the selected rows, columns, and treatments respectively, in order of their selection. We shall, however, use the convention that, for example, if $R = t$, then i^* and i are to be the same index; that is if $R = t$ then the range of i^* is not to be taken as a random ordering of the range of i.

Let $z_{i^*j^*}$ denote the observation coresponding to selected row i^* and selected column j^*. Then all our data are given by observations on the set of t^2 random variables $\{z_{i^*j^*}\}$. For particular values of i^* and j^*, $z_{i^*j^*}$ will be a random variable since the population row and column corresponding to (i^*, j^*) will, in general, be a random event; also, because of choosing a random latin square, the treatment assigned to (i^*j^*) will be a random event. This notation classifies our data according to selected rows and columns.

To classify our data according to treatments we will use the symbol x_{k^*f} to denote the f-th replicate of selected treatment k^*, where f has range $1, 2, \ldots, t$ for each value of k^*. Then of course the particular experimental unit, in the population of RC experimental units, which corresponds to x_{k^*f} depends on a random event.

To write explicit models for the $\{z_{i^*j^*}\}$, and the $\{x_{k^*f}\}$, in terms

of the population parameters, it is helpful and convenient to define some additional random variables which will reflect the design (and randomization procedure) of the experiment. The purposes for which the model is useful are to help our interpretation of the experiment by allowing us to focus on parameters of interest, to enable the unambiguous application in the model of knowledge about the physical situation, to enable the assessment of analytic assumptions in terms of the physical meaning of the assumptions, and to make clear those conditions which must exist if the analysis of the experiment is to be meaningful.

Let $a_i^{i^*}$ = 1, if selected row i* corresponds to row i in the population,
= 0, otherwise;

$\beta_j^{j^*}$ = 1, if selected column j* corresponds to column j in population,
= 0, otherwise;

$\gamma_k^{k^*}$ = 1, if selected treatment k* corresponds to treatment k in the population,
= 0, otherwise;

$\delta_{i^*j^*}^{k^*f}$ = 1, if the f-th replicate of selected treatment k* falls on selected unit (i*, j*)
= 0, otherwise.

$\rho_{i^*j^*}^{k^*} = \sum_f \delta_{i^*j^*}^{k^*f}$.

These quantities are random variables whose joint distribution is specified by the experimental procedure and design. Some properties of these random variables are detailed below. In particular, the rule stated by

Fisher, referred to above, determines the given properties of the δ's and ρ's. In what follows immediately a primed index, say i', is to be taken as unequal to the corresponding unprimed index, say i, unless the contrary is stated.

$$E(a_i^{i*}) = \frac{1}{R}\ ;\quad E(\beta_j^{j*}) = \frac{1}{C}\ ;\quad E(\gamma_k^{k*}) = \frac{1}{T}\ ;$$

$$E(a_i^{i*}\, a_{i'}^{i*'}) = \frac{1}{R(R-1)}\ ;\quad E(\beta_j^{j*}\beta_{j'}^{j*'}) = \frac{1}{C(C-1)};\quad E(\gamma_k^{k*}\gamma_{k'}^{k*'}) = \frac{1}{T(T-1)}\ ;$$

$$P(a_i^{i*}\, a_{i'}^{i*'} = 0) = 1;\quad P(a_i^{i*}\, a_i^{i*'} = 0) = 1;$$

$$E(\delta_{i*j*}^{k*f}) = \frac{1}{t^2}\ ;\quad E(\delta_{i*j*}^{k*f}\ \delta_{i*'j*'}^{k*f'}) = \frac{1}{t^2(t-1)^2}\ ;$$

$$E(\delta_{i*j*}^{k*f}\ \delta_{i*'j*}^{k*'f'}) = \frac{1}{t^3(t-1)}\ ,\ \text{all } f,\ f'\ ;$$

$$P(\delta_{i*j*}^{k*f}\ \delta_{i*j*'}^{k*f'} = 0) = 1;\quad P(\delta_{i*j*}^{k*f}\ \delta_{i*'j*}^{k*f'} = 0) = 1\ ;$$

$$E(\delta_{i*j*}^{k*f}\ \delta_{i*'j*'}^{k*'f'}) = \frac{(t-2)}{t^3(t-1)^2}\ ,\ \text{all } f,\ f'\ ;$$

$$E(\rho_{i*j*}^{k*}) = \frac{1}{t}\ ;\quad E(\rho_{i*j*}^{k*}\rho_{i*'j*}^{k*'}) = \frac{1}{t(t-1)}\ ;$$

$$P(\rho_{i*j*}^{k*}\ \rho_{i*'j*}^{k*} = 0) = 1\ ;$$

$$E(\rho_{i*j*}^{k*}\ \rho_{i*'j*'}^{k*}) = \frac{1}{t(t-1)}\ ;\quad E(\rho_{i*j*}^{k*}\ \rho_{i*'j*'}^{k*'}) = \frac{(t-2)}{t(t-1)^2}\ ;$$

the $\{a_i^{i*}\}$, $\{\beta_j^{j*}\}$, $\{\gamma_k^{k*}\}$, and $\{\rho_{i*j*}^{k*}\}$ are statistically independent; also

$$P(\Sigma_i\, a_i^{i*} = 1) = 1;\ P(\Sigma_{k*}\, \rho_{i*j*}^{k*} = 1);\ P(\Sigma_{i*}\, \rho_{i*j*}^{k*} = 1) = 1.$$

The statistical models are as follows:

$$z_{i^*j^*} = \sum_{ijkk^*} a_i^{i^*} \beta_j^{j^*} \gamma_k^{k^*} \rho_{i^*j^*}^{k^*} y_{ijk}$$

$$= \mu + \sum_{i} a_i^{i^*} r_i + \sum_{j} \beta_j^{j^*} c_j + \sum_{kk^*} \gamma_k^{k^*} \beta_k^{k^*} \rho_{i^*j^*}^{k^*} t_k + \sum_{ikk^*} a_i^{i^*} \gamma_k^{k^*} \rho_{i^*j^*}^{k^*} (rt)_{ik}$$

$$+ \sum_{jkk^*} \beta_j^{j^*} \gamma_k^{k^*} \rho_{i^*j^*}^{k^*} (ct)_{jk} + \sum_{ij} a_i^{i^*} \beta_j^{j^*} e_{ij} + \sum_{ijkk^*} a_i^{i^*} \beta_j^{j^*} \gamma_k^{k^*} \rho_{i^*j^*}^{k^*} \epsilon_{ijk} \; ;$$

$$x_{k^*f} = \sum_{i^*j^*} \delta_{i^*j^*}^{k^*f} z_{i^*j^*} \; .$$

The random variables in these models are the $\left\{ a_i^{i^*} \right\}$, $\left\{ \beta_j^{j^*} \right\}$, $\left\{ \gamma_k^{k^*} \right\}$, $\left\{ \delta_{i^*j^*}^{k^*f} \right\}$, and $\left\{ \rho_{i^*j^*}^{k^*} \right\}$, which take on the values 0 and 1 with probabilities specified by the experimental design, and the ϵ_{ijk}. The advantages of such a statement of the statistical model is that the population parameters appear explicitly and the accomplishments of randomization can be studied objectively through the random variables $\left\{ a_i^{i^*} \right\}$, $\left\{ \beta_j^{j^*} \right\}$, etc.

In what follows we shall use these statistical models in studying the analysis of variance used for the latin square design, and certain simple estimation questions.

5. The analysis of variance

The analysis of variance for a latin square design is detailed in Table 27 in terms of the notation developed above. We use the dot convention to denote means.

An explicit expression for the discrepance sum of squares is

$$D' = \sum_{i^*j^*} (z_{i^*j^*} - z_{i^*.} - z_{.j^*} + z_{..})^2 - \sum_{k^*f} (x_{k^*.} - x_{..})^2 \; .$$

Table 27. Analysis of variance for the latin square design

Due to	d.f.	Sum of Squares	Mean Squares
Rows	(t-1)	$R' = t \sum_{i*} (z_{i*.} - z_{..})^2$	$R^* = R'/(t-1)$
Columns	(t-1)	$C' = t \sum_{j*} (z_{.j*} - z_{..})^2$	$C^* = C'/(t-1)$
Treatments	(t-1)	$T' = t \sum_{k*} (x_{k*.} - x_{..})^2$	$T^* = T'/(t-1)$
Discrepance	(t-1)(t-2)	$D' = G' - R' - C' - T'$	$D^* = D'/(t-1)(t-2)$
Total	t^2-1	$G' = \sum_{i*j*} (z_{i*j*} - z_{..})^2$	

It is purely formal notational difficulty which makes it inconvenient to write D' as a sum of squares. It is of course an always non-negative number.

To aid the interpretation of the analysis of variance we give the expectations of the mean squares using the derived statistical model. The evaluation of these expectations is a purely algebraic task, as would be the evaluation of other moments, once the needed properties of the random variables of the model were specified. Some of the algebra is exemplified below. The expected mean squares are given in Table 28, notational definitions are given following Table 28.

We shall outline the details of finding E(G'), where E denotes the usual expectation operator.

$$G' = \sum_{i^*j^*} (z_{i^*j^*} - z_{..})^2 = \sum_{i^*j^*} z^2_{i^*j^*} - t^2 z^2_{..} = L - t^2 M, \text{ say.}$$

$$E(L) = \sum_{i^*j^*} E(z^2_{i^*j^*})$$

$$= t^2 \left[\mu^2 + \frac{1}{R} \sum_i r_i^2 + \frac{1}{C} \sum_j c_j^2 + \frac{1}{T} \sum_k t_k^2 + \frac{1}{RT} \sum_{ik} (rt)^2_{ik} \right.$$

$$\left. + \frac{1}{CT} \sum_{jk} (ct)^2_{jk} + \frac{1}{RC} \sum_{ij} e^2_{ij} + \frac{1}{RCT} \sum_{ijk} n^2_{ijk} \right] .$$

This result derives from operations such as:

$$\sum_{i^*j^*} E\Big(\sum_i a_i^{i^*} r_i \Big)\Big(\sum_i a_i^{i^*} r_i \Big) = \sum_{i^*j^*} E\Big(\sum a_i^{i^*} r_i^2 + \sum_{i \neq i'} a_i^{i^*} a_{i'}^{i^*} r_i r_{i'} \Big)$$

$$= \frac{t^2}{R} \sum_i r_i^2 , \text{ since } E(a_i^{i^*}) = \frac{1}{R} \text{ and } P(a_i^{i^*} a_{i'}^{i^*} = 0) = 1;$$

and $$\sum_{i^*j^*} E\Big(\sum_i a_i^{i^*} r_i \Big)\Big(\sum_j \beta_j^{j^*} c_j \Big) = \sum_{i^*j^*} E\Big(\sum_{ij} a_i^{i^*} \beta_j^{j^*} r_i c_j \Big)$$

$$= \frac{t^2}{RC} \Big(\sum_i r_i \Big)\Big(\sum_j c_j \Big) = 0, \text{ since } a_i^{i^*} \text{ and } \beta_j^{j^*} \text{ are independent, and}$$

$$\sum_i r_i = \sum_j c_j = 0.$$

Such a procedure is valid for the case where, say, R = t, as we now show. When R = t, then i* and i are the same index; hence with probability 1, $a_1^1 = 1$, $a_1^2 = 0$, $a_3^3 = 1$, etc. In general $a_{i^*}^{i^*} = 1$, $a_{i'}^{i} = 0$. Thus consider

$$\sum_{i^*j^*} E\Big(\sum_i a_i^{i^*} r_i \Big)\Big(\sum_i a_i^{i^*} r_i \Big) = \sum_{i^*j^*} E(r_{i^*})(r_{i^*}), \text{ since } a_{i^*}^{i} = 0 \text{ for } i \neq i^*,$$

$$= t \sum_i r_i^2 , \text{ since i* and i are the same index; and this is of course}$$

a special case of $\frac{t^2}{R} \sum_i r_i^2$ when R = t.

To find E(M) we can proceed as follows:

$$M = \frac{1}{t^4}\Big(\sum_{i^*j^*} z_{i^*j^*}\Big)^2$$

$$= \frac{1}{t^4}\Big(\sum_{i^*j^*} z^2_{i^*j^*} + \sum_{\substack{i^*\neq i^{*\prime} \\ j^*}} z_{i^*j^*}z_{i^{*\prime}j^*} + \sum_{\substack{i^* \\ j^*\neq j^{*\prime}}} z_{i^*j^*}z_{i^*j^{*\prime}}$$

$$+ \sum_{\substack{i^*\neq i^{*\prime} \\ j^*\neq j^{*\prime}}} z_{i^*j^*}z_{i^{*\prime}j^{*\prime}}\Big)$$

$$= \frac{1}{t^4}(L + M_1 + M_2 + M_3), \text{ say.}$$

In evaluating $E(M_3)$ we encounter expressions such as

$$\sum_{\substack{i^*\neq i^{*\prime} \\ j^*\neq j^{*\prime}}} E\Big(\sum_{\substack{i \\ kk^*}} a_i^{i^*} \gamma_k^{k^*} \rho_{i^*j^*}^{k^*}(rt)_{ik}\Big)\Big(\sum_{\substack{i \\ kk^*}} a_i^{i^{*\prime}} \gamma_k^{k^*} \rho_{j^{*\prime}j^{*\prime}}^{k^*}(rt)_{ik}\Big)$$

$$= \sum_{\substack{i^*\neq i^{*\prime} \\ j^*\neq j^{*\prime}}} E\Big(\sum_{\substack{i\neq i' \\ kk^*}} a_i^{i^*} a_{i'}^{i^{*\prime}} \gamma_k^{k^*} \rho_{i^*j^*}^{k^*} \rho_{i^{*\prime}j^{*\prime}}^{k^*}(rt)_{ik}(rt)_{i'k}$$

$$+ \sum_{\substack{i\neq i' \\ k\neq k' \\ k^*\neq k^{*\prime}}} a_i^{i^*} a_{i'}^{i^{*\prime}} \gamma_k^{k^*} \gamma_{k'}^{k^{*\prime}} \rho_{i^*j^*}^{k^*} \rho_{i^{*\prime}j^{*\prime}}^{k^{*\prime}}(rt)_{ik}(rt)_{i'k'}\Big)$$

$$= t^2(t-1)^2 \Big(\frac{-t}{R(R-1)Tt(t-1)} + \frac{t(t-1)(t-2)}{R(R-1)T(T-1)t(t-1)^2}\Big) \sum_{ik} (rt)^2_{ik}$$

$$= \frac{-t^2(t-1)(T-t+1)}{RT(R-1)T-1)} \sum_{ik} (rt)^2_{ik} .$$

As an example of what happens with cross product terms in M_3,

$$\sum_{\substack{i^* \neq i^{*'} \\ j^* \neq j^{*'}}} E\Big(\sum_i a_i^{i^*} r_i\Big)\Big(\sum_{ij} a_i^{i^{*'}} \beta_j^{j^{*'}} e_{ij}\Big)$$

$$= \sum_{\substack{i^* \neq i^{*'} \\ j^* \neq j^{*'}}} E\Big(\sum_{\substack{i \neq i' \\ j}} a_i^{i^*} a_{i'}^{i^{*'}} \beta_j^{j^{*'}} r_i e_{i'j}\Big), \text{ since } P(a_i^{i^*} a_i^{i^{*'}} = 0) = 1$$

$$= \frac{t^2(t-1)^2}{R(r-1)C} \sum_{i \neq i'} r_i \Big(\sum_j e_{i'j}\Big)$$

$$= 0, \text{ since } \sum_j e_{i'j} = 0.$$

It will be found that

$$E(G') = (t-1)\Big[\frac{t}{R-1}\sum_i r_i^2 + \frac{t}{C-1}\sum_j c_j^2 + \frac{t}{T-1}\sum_k t_k^2$$

$$+ (t + 1 - \frac{t}{T} - \frac{t}{R}) \frac{1}{(R-1)(T-1)} \sum_{ik} (rt)_{ik}^2$$

$$+ (t + 1 - \frac{t}{T} - \frac{t}{C}) \frac{1}{(C-1)(T-1)} \sum_{jk} (ct)_{jk}^2 + (t + 1 - \frac{t}{R} - \frac{t}{C})$$

$$\frac{1}{(R-1)(C-1)} \sum_{ij} e_{ij}^2$$

$$+ (t + 1 - \frac{t+1}{T} - \frac{t+1}{R} - \frac{t+1}{C} + \frac{t}{RC} + \frac{t}{RT} + \frac{t}{CT}) \frac{1}{(R-1)(C-1)(T-1)} \sum_{ijk} n_{ijk}^2$$

$$+ (t+1)\sigma^2\Big].$$

The expectations of R', C', T' and D' may be obtained from E(L), $E(M_1)$, $E(M_2)$ and $E(M_3)$. The results on expected mean squares are given in Table 28.

Table 28. Expected mean squares for latin square design

Mean Square	Expectation of Mean Square
R*	$\sigma^2 + (\pi + \frac{t}{CT})\,\sigma_n^2 + \frac{(C-t)}{C}\sigma_e^2 + \sigma_{ct}^2 + \frac{(T-t)}{T}\sigma_{rt}^2 + t\sigma_r^2$
C*	$\sigma^2 + (\pi + \frac{t}{RT})\sigma_n^2 + \frac{(R-t)}{R}\sigma_e^2 + \frac{(T-t)}{T}\,\sigma_{ct}^2 + \sigma_{rt}^2 + t\sigma_c^2$
T*	$\sigma^2 + (\pi + \frac{t}{RC})\,\sigma_n^2 + \sigma_e^2 + \frac{(C-t)}{C}\sigma_{ct}^2 + \frac{(R-t)}{R}\sigma_{rt}^2 + t\sigma_t^2$
D*	$\sigma^2 + \pi\sigma_n^2 + \sigma_e^2 + \sigma_{ct}^2 + \sigma_{rt}^2$

The definitions of the various symbols used in Table 28 are as follows:

$$\sigma_r^2 = \frac{1}{R-1}\sum_i r_i^2\,;\quad \sigma_c^2 = \frac{1}{C-1}\sum_j c_j^2\,;\quad \sigma_t^2 = \frac{1}{T-1}\sum_k t_k^2\,;$$

$$\sigma_{rt}^2 = \frac{1}{(R-1)(T-1)}\sum_{ik}(rt)_{ik}^2\,;\quad \sigma_{ct}^2 = \frac{1}{(C-1)(T-1)}\sum_{jk}(ct)_{jk}^2\,;$$

$$\sigma_e^2 = \frac{1}{(R-1)(C-1)}\sum_{ij} e_{ij}^2\ (=\sigma_{rc}^2);\quad \sigma_n^2 = \frac{1}{(R-1)(C-1)(T-1)}\sum_{ijk} n_{ijk}^2\ (=\sigma_{rct}^2);$$

$$\sigma^2 = E(\epsilon_{ijk}^2)\,;\quad \pi = (1 - \frac{1}{R} - \frac{1}{C} - \frac{1}{T})\,.$$

It is interesting to note the symmetry of the expected mean squares for rows, columns and treatments. This is in contrast to the asymmetry between blocks and treatments exhibited in similar expressions for the randomized block design.

In general, it appears that the "error mean square" is "too small" on the average in that the coefficient of σ_n^2 in E(D*) is less than the

corresponding coefficients for E(R*), E(C*), E(T*). Also, in general, the existence of row-treatment or column-treatment interactions will confuse the interpretation of the analysis of variance in that the coefficients of, for example, σ^2_{rt} in E(T*) may be quite different from the corresponding quantity in E(D*). Hence the estimation of σ^2_t or the assessment of the experiment to give adequate estimates of the relative magnitudes of treatment effects may be impossible or quite ambiguous. Evidently interactions of treatments with rows or columns do not enter with the same weights, on the average, in the treatments and discrepance mean squares.

The results of Table 28 may be put in a more symmetric form which parallels entirely the pattern of results for the completely randomized, randomized blocks, and split-plot designs. This form is detailed in Table 29. We recall that σ^2_e reflects $\mathcal{R}$ x $\mathcal{C}$ interactions and will be written as σ^2_{rc}, while σ^2_n reflects $\mathcal{R}$ x $\mathcal{C}$ x $\mathcal{T}$ interactions and will be written σ^2_{rct}. The definitions used in Table 29 are as follows:

$$\Sigma_r = \sigma^2_r - \frac{1}{C}\sigma^2_{rc} - \frac{1}{T}\sigma^2_{rt} + \frac{1}{CT}\sigma^2_{rct} ,$$

$$\Sigma_c = \sigma^2_c - \frac{1}{R}\sigma^2_{rc} - \frac{1}{T}\sigma^2_{ct} + \frac{1}{RT}\sigma^2_{rct} ,$$

$$\Sigma_t = \sigma^2_t - \frac{1}{R}\sigma^2_{rt} - \frac{1}{C}\sigma^2_{ct} + \frac{1}{RC}\sigma^2_{rct} ,$$

$$\Sigma_{rc} = \sigma^2_{rc} - \frac{1}{T}\sigma^2_{rct} ,$$

$$\Sigma_{rt} = \sigma^2_{rt} - \frac{1}{C}\sigma^2_{rct} ,$$

$$\Sigma_{ct} = \sigma^2_{ct} - \frac{1}{R}\sigma^2_{rct} \quad ,$$

$$\Sigma_{rct} = \sigma^2_{rct} \quad ,$$

$$\Sigma_0 = \sigma^2 + \Sigma_{rct} + \Sigma_{ct} + \Sigma_{rt} + \Sigma_{rc} \quad .$$

Table 29. Expected mean squares for latin square design (Alternate form)

Mean Square	Expected Mean Square
R*	$t\,\Sigma_r + \Sigma_0$
C*	$t\,\Sigma_c + \Sigma_0$
T*	$t\,\Sigma_t + \Sigma_0$
D*	$\Sigma_0 = \sigma^2 + \Sigma_{rct} + \Sigma_{rt} + \Sigma_{ct} + \Sigma_{rc}$

It is evident from Table 29 just what are the linear combinations of components of variation which are always estimable regardless of assumptions or relation of population and sample sizes.

The form of the inverse relationship of the σ^2)s to the Σ's is exemplified by

$$\sigma^2_r = \Sigma_r + \frac{1}{C}\Sigma_{rc} + \frac{1}{T}\Sigma_{rt} + \frac{1}{CT}\Sigma_{rct} \quad .$$

Table 30 details the results for some special cases of, perhaps, particular interest. We note that, in the general case where interactions may be present in our data, the latin square "error mean square"

Table 30. Some special cases

Case	I. $R = C = T = t$	II. $R \gg t$; $C \gg t$; $T = t$	III. $R = C = t$; $T \gg t$	IV. $R \gg t$; $C \gg t$; $T \gg t$
M.S.	E.M.S.	E.M.S.	E.M.S.	E.M.S.
R*	$\sigma^2 + (1 - \frac{2}{t})\ \sigma_n^2 + \sigma_{ct}^2 + t\sigma_r^2$	$\sigma^2 + (1 - \frac{1}{t})\ \sigma_n^2 + \sigma_e^2 + \sigma_{ct}^2 + t\sigma_r^2$	$\sigma^2 + (1 - \frac{2}{t})\ \sigma_n^2 + \sigma_{ct}^2 + \sigma_{rt}^2 + t\sigma_r^2$	$\sigma^2 + \sigma_n^2 + \sigma_e^2 + \sigma_{ct}^2 + \sigma_{rt}^2 + t\sigma_r^2$
C*	$\sigma^2 + (1 - \frac{2}{t})\ \sigma_n^2 + \sigma_{rt}^2 + t\sigma_c^2$	$\sigma^2 + (1 - \frac{1}{t})\ \sigma_n^2 + \sigma_e^2 + \sigma_{rt}^2 + t\sigma_e^2$	$\sigma^2 + (1 - \frac{2}{t})\ \sigma_n^2 + \sigma_{ct}^2 + \sigma_{rt}^2 + \sigma_c^2$	$\sigma^2 + \sigma_n^2 + \sigma_e^2 + \sigma_{ct}^2 + \sigma_{rt}^2 + t\sigma_c^2$
T*	$\sigma^2 + (1 - \frac{2}{t})\ \sigma_n^2 + \sigma_e^2 + t\sigma_t^2$	$\sigma^2 + (1 - \frac{1}{t})\ \sigma_n^2 + \sigma_e^2 + \sigma_{ct}^2 + \sigma_{rt}^2 + t\sigma_t^2$	$\sigma^2 + (1 - \frac{1}{t})\ \sigma_n^2 + \sigma_e^2 + t\sigma_t^2$	$\sigma^2 + \sigma_n^2 + \sigma_e^2 + \sigma_{ct}^2 + \sigma_{rt}^2 + t\sigma_t^2$
D*	$\sigma^2 + (1 - \frac{3}{t})\ \sigma_n^2 + \sigma_e^2 + \sigma_{ct}^2 + \sigma_{rt}^2$	$\sigma^2 + (1 - \frac{1}{t})\ \sigma_n^2 + \sigma_e^2 + \sigma_{ct}^2 + \sigma_{rt}^2$	$\sigma^2 + (1 - \frac{2}{t})\ \sigma_n^2 + \sigma_e^2 + \sigma_{ct}^2 + \sigma_{rt}^2$	$\sigma^2 + \sigma_n^2 + \sigma_e^2 + \sigma_{ct}^2 + \sigma_{rt}^2$

tends to overestimate the appropriate error for evaluting the treatment effects observed in all cases except Case IV, which in the usual terminology would be "rows, columns and treatments all random".

It should be noted that Cases II and IV do not correspond to choosing t^2 units at random from a large population of RC experimental units and then applying treatments so as to have a latin square configuration. The sampling of experimental units employed in the situation under consideration is more restrictive in that only t rows and t columns can be represented. Hence it is generally not true that experiments carried out in a latin square design can be considered as corresponding to Cases II or IV. It is however easy to conceive of experimental situations which might be handled so that $R > t$, $C > t$, and belong to Cases II or IV of Table 30 in particular instances.

It is relevant to note that even if σ_t^2 is zero the treatment mean square can be greater than the error mean square, if interactions of treatments with rows and columns are absent but second order interactions of treatments with rows and columns exist. It is more important to note that even if, as indicated earlier, there may be some reason to take the n_{ijk} to be zero, the existence of row-treatment and column-treatment interactions may cause the treatment mean square to be less, on the average, than the error mean square even if σ_t^2 is greater than zero. It should perhaps be mentioned in recapitulation that σ_t^2 measures the variation among the mean responses of the different treatments over the whole experimental material.

It would appear from these considerations that the latin square is a design of limited applicability. From randomization test considerations the 3 x 3 and 4 x 4 squares are not satisfactory (see Kempthorne (1952a, 1952b)) and with a large number of rows and columns it is likely that row-treatment and column-treatment interactions will be non-negligible. It would be rare in experimentation that the response from a treatment does not depend on the level of response under some standard treatment. The latin square is effective and desirable only under the circumstances that there are sizeable row and column "effects" and this is precisely when one would expect row-treatment and column-treatment interactions. The design has the further drawback that the possible existence of such interactions cannot be checked directly.

6. Estimation

It is clear from Table 28 or 29 that, with non-additivity of rows and treatments and columns and treatments, it is in general not possible to obtain reasonable estimates of the component σ_t^2. Only in the case where $R \gg t$, $C \gg t$ is it true that $E(T^* - D^*) = \sigma_t^2$, when row-treatment or column-treatment interactions exist.

In order to estimate differences of treatment effects, say $(t_k - t_{k'})$, for the case where the treatments are a fixed set, i.e. $T = t$, we might use $(x_{k.} - x_{k'.})$. For the case $T = t$ our statistical model for $x_{k.}$ becomes

$$x_{k.} = \mu + t_k + \frac{1}{t}\left[\sum_{ii^*} a_i^{i^*}(r_i + (rt)_{ik}) + \sum_{jj^*} \beta_j^{j^*}(c_j + (ct)_{jk}) + \sum_{ii^*jj^*} a_i^{i^*}\beta_j^{j^*}\rho_{i^*j^*}^{k}(e_{ij} + n_{ijk} + \epsilon_{ijk})\right] .$$

Thus $E(x_{k.} - x_{k'.}) = t_k - t_{k'}$.

In general, unbiased estimates of linear functions such as $\sum_k d_k t_k$ will be given by $\sum_k d_k \hat{t}_k$, where $\hat{t}_k = x_{k.} - x_{..}$.

The variance of the estimate $x_{k.} - x_{k}$ is $Var(x_{k.} - x_{k'.}) = E\left[(x_{k.} - x_{k'.}) - (t_k - t_{k'})\right]^2$, which can be worked out using the statistical model to give

$$\frac{2}{t}\Bigg[\sigma^2 + \sigma_e^2 + \frac{R-t}{R}\,\frac{1}{(R-1)}\,\sum_i\left((rt)^2_{ik} + (rt)^2_{ik'} - 2(rt)_{ik}(rt)_{ik'}\right)$$

$$+\,\frac{C-t}{C}\,\frac{1}{C-1}\,\sum_j (ct)^2_{jk} + (ct)^2_{jk'} - 2(ct)_{jk}(ct)_{jk'})$$

$$+\left(1 - \frac{1}{C} - \frac{1}{R} + \frac{t}{RC}\right)\frac{1}{(R-1)(C-1)}\,\sum_{ij}(n^2_{ijk} + n^2_{ijk'})$$

$$+\left(1 + \frac{1}{C} + \frac{1}{R} - \frac{t}{RC}\right)\frac{1}{(R-1)(C-1)}\,\sum_{ij}(n_{ijk} + n_{ijk'})\,e_{ij}$$

$$+\,2\left(\frac{1}{R} + \frac{1}{C} - \frac{t}{RC}\right)\frac{1}{(R-1)(C-1)}\,\sum_{ij} n_{ijk}\,n_{ijk'}\Bigg] \quad .$$

Of course it is impossible to estimate this quantity, but a useful related quantity is the average variance for all possible treatment differences. This is

$$\bar{V}\,(\text{diff.}) = \frac{1}{t(t-1)}\,\sum_{k\neq k'} Var(x_{k.} - x_{k'.})$$

$$= \frac{2}{t}\Bigg[\sigma^2 + \sigma_e^2 + \left(1 - \frac{1}{t} - \frac{1}{C} - \frac{1}{R} + \frac{t}{RC}\right)\sigma_n^2$$

$$+\,\frac{(R-t)}{R}\,\sigma^2_{rt} + \frac{(C-t)}{C}\,\sigma^2_{ct}\Bigg] \quad .$$

The coefficient of σ_n^2 in the above expression is equal to $(\pi + \frac{t}{RC})$ in the notation of Table 28. The conditions under which D* can be used to give a reasonable estimate of $\bar{V}$(diff.) may be determined by comparing with E(D*) given in Table 28, or from Table 30.

When R = C = t, $\bar{V}$(diff.) becomes

$$\frac{2}{t}\left[\sigma^2 + \sigma_e^2 + \sigma_n^2\right] \quad .$$

A main objective of using the latin square design is that the variance of estimates of differences such as $(t_k - t_{k'})$ will be made small by making the comparisons "within rows" and also "within columns". While this may in fact be accomplished by the design we see that, in the presence of non-additivities of treatments with rows and columns, the latin square analysis of variance may seriously overestimate this variance.

D. The Split Plot Design

From some points of view the split plot design is a generalization of the randomized block design. However, it introduces several additional features. In the simplest form of the design the experimental units have a three-fold hierarchal structure, say units within blocks within sources, while the treatments have a two-factor structure, say $\mathcal{G}$ and $\mathcal{H}$. However levels of $\mathcal{G}$ are applied to blocks while levels of $\mathcal{H}$ are applied to units. Thus certain treatment comparisons are confounded with units within blocks while others are confounded with blocks. The details of the design are made precise in the next section.

For the simplest split plot structure we shall describe in detail the development of a population model and a statistical model for a general experimental situation and general design conditions, having certain properties of balance; and shall detail some results in connection with the analysis of variance and estimation questions. Attention will then be given to extension of some of the results to more complex split plot type structures.

1. The experimental situation and design

Consider two sets of treatments $\mathcal{G}$ and $\mathcal{H}$. We shall refer to these sets as factors. Let factor $\mathcal{G}$ have G levels and factor $\mathcal{H}$ have H levels.

Suppose we have SBP experimental units which are structured into SB "blocks" each containing P units; the blocks in turn are structured into S "sources" each containing B blocks. (For example, sources might be various oil wells; a block might be a day's production; a unit might be a barrel of oil.) (The notation and nomenclature are intended to be suggestive but not restrictive.)

To study this experimental situation we proceed according to the following design:

(i) Select g levels of $\mathcal{G}$ from G at random.

(ii) Select h levels of $\mathcal{H}$ from H at random.

(iii) Select s sources from S at random.

(iv) From each selected source, select dg blocks at random.

(v) From each selected block, select rh units at random.

(vi) Apply selected combinations of levels of $\mathcal{G}$ and $\mathcal{H}$ to the selected

experimental units at random but so that (referring to selected entities) each level of $\mathcal{H}$ falls on r units in every block, and each level of $\mathcal{G}$ falls on d blocks in every source.

The correspondence to the usual terminology of the split-plot design is as follows:

$\mathcal{G}$ represents whole plot treatments.

$\mathcal{H}$ represents split-plot treatments.

Sources are often referred to as replicates.

Blocks within sources are the whole plots.

Units without blocks are the split plots.

The design has been generalized to allow for replication of factor levels within sources and/or within blocks, though in most applications of this design d = r = 1.

Special cases of the situation and design described will revolve around the relationships of population sizes to experiment sizes; for example, whole-plot treatments fixed corresponds to $G = g$, which we would refer to as "the factor $\mathcal{G}$ fixed".

Our index (subscript) notation shall be as follows:

$i = 1, 2, \ldots, S$ denotes the source in the population of sources;

$j = 1, 2, \ldots, B$ denotes the block within a given source;

$k = 1, 2, \ldots, P$ denotes the unit within a given block;

$m = 1, 2, \ldots, G$ denotes the levels of factor $\mathcal{G}$ in the population of levels of $\mathcal{G}$;

$n = 1, 2, \ldots, H$ denotes the levels of factor $\mathcal{H}$ in the population of levels of $\mathcal{H}$.

$i^* = 1, 2, \ldots s$ denotes the selected sources, in order of their selection;

$m^* = 1, 2, \ldots, g$ denotes the selected levels of $\mathcal{G}$, in order of their selection;

$n^* = 1, 2, \ldots, h$ denotes the selected levels of $\mathcal{H}$, in order of their selection.

$f = 1, 2, \ldots, d$ denotes replicates of selected levels of $\mathcal{G}$ in a given source.

$e = 1, 2, \ldots, r$ denotes replicates of selected levels of $\mathcal{H}$ in a given block.

As usual, we make the convention that if $S = s$ then i^* and i are the same index; if $G = g$ then m^* and m are the same index; if $H = h$ then n^* and n are the same index.

2. The population model

Let Y_{ijkmn} represent the true (unknown) response if level m of $\mathcal{G}$ and level n of $\mathcal{H}$ were applied to the k-th unit of the j-th block of source i. We take as our population of interest the set of SBPGH numbers $\{Y_{ijkmn}\}$. This set of (conceptual) numbers represents the conceivable totality of experimental knowledge which we might acquire in this situation.*

We can now write the identity

$$Y_{ijkmn} = \mu + s_i + g_m + (sg)_{im} + b_{ij} + (bg)_{ijm} + h_n$$

*A fundamental assumption here as elsewhere in the thesis is that the response from treatment (mn) on unit (ijk) depends only on the particular treatment and unit.

$$+ (sh)_{in} + (gh)_{mn} + (sgh)_{imn} + (bh)_{ijn}$$

$$+ (bgh)_{ijmn} + p_{ijk} + q_{ijkmn} \ ,$$

where:

$\mu = Y_{.....}$ is the (conceptual) overall mean response from all combinations on all experimental units.

$s_i = (Y_{i....} - \mu)$ is the difference between the overall average response on source i, and μ; we call s_i the effect of source i.

Similarly, $g_m = (Y_{...m.} - \mu)$ is the effect of the m-th level of $\mathcal{G}$.

Similarly, $h_n = (Y_{....n} - \mu)$ is the effect of the n-th level of $\mathcal{H}$.

$(sg)_{im} = (Y_{i..m.} - Y_{i...} - Y_{...m.} + Y_{.....})$ is the difference between the effect of the m-th level of $\mathcal{G}$ on the i-th source and its effect over all the sources; we call $(sg)_{im}$ the interaction of the i-th source and the m-th level of $\mathcal{G}$.

Similarly, $(sh)_{in} = (Y_{i...n} - Y_{i....} - Y_{....n} + Y_{.....})$ is the interaction of the i-th source and the n-th level of $\mathcal{H}$.

Similarly, $(gh)_{mn} = (Y_{...mn} - Y_{...m.} - Y_{....n} + Y_{.....})$ is the interaction of the m-th level of $\mathcal{G}$ and the n-th level of $\mathcal{H}$.

Similarly, $(bg)_{ijm} = (Y_{ij.m.} - Y_{ij...} - Y_{i..m.} + Y_{i....})$ is the within i-th source interaction of the j-th block of source i with level m of $\mathcal{G}$.

Similarly, $(bh)_{ijn} = (Y_{ij..n} - Y_{ij...} - Y_{i...n} + Y_{i....})$ is the within i-th source interaction of the j-th block of source i with level n of $\mathcal{H}$.

$b_{ij} = (Y_{ij...} - Y_{i....})$ is the difference between the overall average response on block j of source i and the average over all blocks of source i.

Hence b_{ij} measures the within i-th source deviation of block j of source i from the average of the source with respect to an "average treatment". We shall refer to b_{ij} as the within i-th source effect of block j, but we note that it behaves as an additive unit error with respect to factor $\mathcal{G}$.

$$(sgh)_{imn} = (Y_{i..mn} - Y_{i..m.} - Y_{i...n} + Y_{i....} - (gh)_{mn})$$

is the difference between the interaction of the m-th level of $\mathcal{G}$ and the n-th level of $\mathcal{H}$ on source i and their interaction over all the sources. We call $(sgh)_{imn}$ the interaction of the i-th source with the m-th level of $\mathcal{G}$ and the n-th level of $\mathcal{H}$.

Similarly,

$$(bgh)_{ijmn} = (Y_{ij.mn} - Y_{ij.m.} - Y_{ij..n} + Y_{ij...} - Y_{i..mn} + Y_{i..m.} + Y_{i...n} - Y_{i....})$$

is the within i-th source interaction of block j of source i with the m-th level of $\mathcal{G}$ and the n-th level of $\mathcal{H}$.

$p_{ijk} = (Y_{ijk..} - Y_{ij...})$ is the deviation of the average response on unit k of block j in source i from the average over all units in that block. We shall call p_{ijk} the within j-th block of source i additive unit error of unit k.

$q_{ijkmn} = (Y_{ijkmn} - Y_{ijk..} - Y_{ij.mn} + Y_{ij...})$ represents a within j-th block of source i interaction of unit k with the treatment combination consisting of the m-th level of $\mathcal{G}$ and the n-th level of $\mathcal{H}$.

We shall, as before, assume that our conceptual observable is a random variable y_{ijkmn} such that

$$y_{ijkmn} = Y_{ijkmn} + \epsilon_{ijkmn}$$

where the ϵ_{ijkmn} are uncorrelated with zero means and constant variance σ^2. The ϵ_{ijkmn} are considered as representing variability of techniques of measurement and treatment application.

3. The statistical model

The design of the experiment is such that we obtain a random, within restrictions, sample from the set of conceptual observables $\{y_{ijkmn}\}$.

Let $x_{i*m*fn*e}$ denote the e-th replicate of selected level n* of $\mathcal{H}$ in the block to which is applied the f-th replicate of selected level m* of $\mathcal{G}$ in the i*-th selected source. To write an explicit model for the $x_{i*m*fn*e}$ in terms of the population parameters we define the following quantities:

Let α_i^{i*} = 1 if selected source i* corresponds to source i in the population,

= 0 otherwise.

Let γ_m^{m*} = 1 if selected level m* of $\mathcal{G}$ corresponds to the m-th level of $\mathcal{G}$ in the population,

= 0 otherwise.

Let δ_n^{n*} = 1 if selected level n* of $\mathcal{H}$ corresponds to the n-th level of $\mathcal{H}$ in the population,

= 0 otherwise.

Let β_{i*j}^{i*m*f} = 1 if the f-th replicate of selected level m* of $\mathcal{G}$ on selected source i* appears on block j of that source,

= 0 otherwise.

Let $\rho^{i*m*fn*e}_{i*m*fk}$ = 1 if the e-th replicate of selected level n* of $\mathcal{H}$ appears on unit k of the selected block within selected source i* on which the f-th replicate of selected level m* of $\mathcal{G}$ fell,

= 0 otherwise.

These quantities are random variables because of random methods of selection and allocation. Their distributional properties derive from the experimental design. For example:

The sets of α's, γ's, δ's, β's and ρ's are groupwise independent.

$$P(\alpha_i^{i*} = 1) = \frac{1}{S}$$

$$P(\alpha_i^{i*}\alpha_{i'}^{i*'} = 1) = \frac{1}{S(S-1)}, \quad i \neq i', \quad i* \neq i*' .$$

$$P(\alpha_i^{i*}\alpha_i^{i*'} = 0) = 1, \quad i* \neq i*' .$$

Any element of the set $\{\beta^{i*m*f}_{i*j}, \text{ i* fixed}\}$ is statistically independent of any element of the set $\{\beta^{i*'m*f}_{i*'j}, \text{ i*' fixed}\}$, for $i* \neq i*'$.

$$P(\beta^{i*m*f}_{i*j} = 1) = \frac{1}{B} .$$

$$P(\beta^{i*m*f}_{i*j}\beta^{i*m*'f'}_{i*j'} = 1) = \frac{1}{B(B-1)}, \quad j \neq j', \ (m*f) \neq (m*'f') .$$

$$P(\rho^{i*m*fn*e}_{i*m*fk}\rho^{i*m*fn*'e'}_{i*m*fk'} = 1) = \frac{1}{P(P-1)}, \quad k \neq k', \ (n*e) \neq (n*'e') .$$

etc.

We can now write the following statistical model for the observations:

$$x_{i*m*fn*e} = \sum_{ijkmn} \alpha_i^{i*}\gamma_m^{m*}\delta_n^{n*}\beta^{i*m*f}_{i*j}\rho^{i*m*fn*e}_{i*m*fk}(Y_{ijkmn} + \epsilon_{ijkmn})$$

$$= \mu + \sum_i a_i^{i*} s_i + \sum_m \gamma_m^{m*} g_m + \sum_n \delta_n^{n*} h_n + \sum_{im} a_i^{i*} \gamma_m^{m*} (sg)_{im}$$

$$+ \sum_{in} a_i^{i*} \delta_n^{n*} (sh)_{in} + \sum_{mn} \gamma_m^{m*} \delta_n^{n*} (gh)_{mn} + \sum_{imn} a_i^{i*} \gamma_m^{m*} \delta_n^{n*} (sgh)_{imn}$$

$$+ \sum_{ij} a_i^{i*} \beta_{i*j}^{i*m*f} b_{ij} + \sum_{ijm} a_i^{i*} \gamma_m^{m*} \beta_{i*j}^{i*m*f} (bg)_{ijm} + \sum_{ijn} a_i^{i*} \delta_n^{n*} \beta_{i*j}^{i*m*f} (bh)_{ijn}$$

$$+ \sum_{ijmn} a_i^{i*} \gamma_m^{m*} \delta_n^{n*} \beta_{i*j}^{i*m*f} (bgh)_{ijmn} + \sum_{ijk} a_i^{i*} \beta_{i*j}^{i*m*f} \rho_{i*m*fk}^{i*m*fn*e} p_{ijk}$$

$$+ \sum_{ijkmn} a_i^{i*} \gamma_m^{m*} \delta_n^{n*} \beta_{i*j}^{i*m*f} \rho_{i*m*fk}^{i*m*fn*e} q_{ijkmn} + \epsilon_{i*m*fn*e}^{*}$$

where

$$\epsilon_{i*m*fn*e}^{*} = \sum_{ijkmn} a_i^{i*} \gamma_m^{m*} \delta_n^{n*} \beta_{i*j}^{i*m*f} \rho_{i*m*fk}^{i*m*fn*e} \epsilon_{ijkmn} .$$

Under the conditions $S > s$, $G > g$, $H > h$, it is easy to see from the above model that

$$E(x_{i*m*fn*e}) = \mu .$$

If $G = g$, $H = h$, while $S > s$, then by our convention m* and m become the same index and n* and n become the same index and the model becomes

$$x_{i*mfne} = \mu + \sum_i a_i^{i*} s_i + g_m + h_n + \sum_i a_i^{i*} (sg)_{im} + \sum_i a_i^{i*} (sh)_{in}$$

$$+ (gh)_{mn} + \sum_i a_i^{i*} (sgh)_{imn} + \sum_{ij} a_i^{i*} \beta_{i*j}^{i*mf} b_{ij}$$

$$+ \sum_{ij} a_i^{i*} \beta_{i*j}^{i*mf} (bg)_{ijm} + \sum_{ij} a_i^{i*} \beta_{i*j}^{i*mf} (bh)_{ijn} + \sum_{ij} a_i^{i*} \beta_{i*j}^{i*mf} (bgh)_{ijmn}$$

$$+ \sum_{ijk} a_i^{i*} \beta_{i*j}^{i*mf} \rho_{i*mfk}^{i*mfne} (q_{ijkmn} + \epsilon_{ijkmn}) .$$

Under this model, it can be seen that

$$E(x_{i*mfne}) = \mu + g_m + h_n + (gh)_{mn}$$

$$E(x_{.m...}) = \mu + g_m$$

$$E(x_{...n.}) = \mu + h_n$$

$$E(x_{.m.n.} - x_{.m...} - x_{...n.} + x_{.....}) = (gh)_{mn} \quad .$$

The above statistical model summarizes within itself all relevant aspects of the design employed. The only assumptions involved in writing it are (i) that the true response from treatment (mn) on unit (ijk) depends only on the unit and the treatment and not on what happens to other units and other treatments; (ii) those assumptions involving the structure of the technical errors; and (iii) that we are able to "sample at random" and to "randomize".

4. The analysis of variance

The analysis of variance for this design is detailed in Table 31. The mean squares for the various lines in the table will be denoted by replacing the ' by a * in the sum of squares notation, as for example

$$S^* = \frac{1}{(s-1)} S' \quad .$$

Of course the line for $\mathcal{B}$ will not exist unless $d > 1$, while the lines for $\mathcal{B} \times \mathcal{H}$ and $\mathcal{P}$ will not exist unless $r > 1$.

Table 31. Analysis of variance, split plot design.

Due to	d. f.	Sum of Squares
Sources ($\mathcal{S}$)	(s-1)	$S' = dgrh \sum_{i*} (x_{i*....} - x_{.....})^2$
$\mathcal{G}$	(g-1)	$G' = sdrh \sum_{m*} (x_{.m*...} - x_{.....})^2$
$\mathcal{H}$	(h-1)	$H' = sdgr \sum_{n*} (x_{...n*.} - x_{.....})^2$
$\mathcal{S} \times \mathcal{G}$	(s-1)(g-1)	$I'_{SG} = drh \sum_{i*m*} (x_{i*m*...} - x_{i*....} - x_{.m*...} + x_{.....})^2$
$\mathcal{S} \times \mathcal{H}$	(s-1)(h-1)	$I'_{SH} = dgr \sum_{i*n*} (x_{i*..n*.} - x_{i*....} - x_{...n*} + x_{.....})^2$
$\mathcal{G} \times \mathcal{H}$	(g-1)(h-1)	$I'_{GH} = sdr \sum_{m*n*} (x_{.m*.n*.} - x_{.m*...} - x_{...n*.} + x_{.....})^2$
$\mathcal{S} \times \mathcal{G} \times \mathcal{H}$	(s-1)(g-1)(h-1)	$I'_{SGH} = dr \sum_{i*m*n*} (x_{i*m*.n*.} - x_{i*m*...} + x_{i*....} - x_{.m*.n*.} + x_{.m*...} + x_{...n*.} - x_{.....})^2$
Blocks within $\mathcal{S}$ and $\mathcal{G}$ ($\mathcal{B}$)	sg(d-1)	$B' = rh \sum_{i*m*f} (x_{i*m*f..} - x_{i*m*...})^2$
$\mathcal{B} \times \mathcal{H}$	sg(d-1)(h-1)	$I'_{BH} = r \sum_{i*m*fn*} (x_{i*m*fn*.} - x_{i*m*.n*.} - x_{i*m*f..} + x_{i*m*...})^2$

Table 31 (Continued)

Due to	d.f.	Sum of Squares
Units within $\mathcal{B}$ and $\mathcal{H}$ ($\mathcal{P}$)	$sdgh(r-1)$	$P' = \sum_{i*m*fn*e} (x_{i*m*fn*e} - x_{i*m*fn*.})^2$

The following notational definitions are used in the succeeding sections:

$$\sigma_s^2 = \frac{1}{(S-1)} \sum_i s_i^2 ; \quad \sigma_g^2 = \frac{1}{(G-1)} \sum_m g_m^2 ;$$

$$\sigma_h^2 = \frac{1}{(H-1)} \sum_n h_n^2 ; \quad \sigma_{sg}^2 = \frac{1}{(S-1)(G-1)} \sum_{im} (sg)_{im}^2 ;$$

$$\sigma_{sn}^2 = \frac{1}{(S-1)(H-1)} \sum_{in} (sh)_{in}^2 ; \quad \sigma_{gh}^2 = \frac{1}{(G-1)(H-1)} \sum_{mn} (gh)_{mn}^2 ;$$

$$\sigma_{sgh}^2 = \frac{1}{(S-1)(G-1)(H-1)} \sum_{imn} (sgh)_{imn}^2 ; \quad \sigma_b^2 = \frac{1}{S(B-1)} \sum_{ij} b_{ij}^2 ;$$

$$\sigma_p^2 = \frac{1}{SB(P-1)} \sum_{ijk} p_{ijk}^2 ; \quad \sigma^2 = E(\epsilon_{i*m*n*}^*)^2 ;$$

$$Q_{gb}^2 = \frac{1}{(G-1)S(B-1)} \sum_{ijm} (bg)_{ijm}^2 ; \quad Q_{hb}^2 = \frac{1}{(H-1)S(B-1)} \sum_{ijn} (bh)_{ijn}^2 ;$$

$$Q_{ghb}^2 = \frac{1}{(G-1)(H-1)(S(B-1)} \sum_{ijmn} (bgh)_{ijmn}^2 ;$$

$$Q_{gp}^2 = \frac{1}{(G-1)SB(P-1)} \sum_{ijkm} q_{ijkm.}^2 ;$$

$$Q_{hp}^2 = \frac{1}{(H-1)SB(P-1)} \sum_{ijkn} q_{ijk.n}^2 ;$$

$$\Omega^2_{ghp} = \frac{1}{(G-1)(H-1)SB(P-1)} \sum_{ijkmn} (q_{ijkmn} - q_{ijkm.} - q_{ijk.n})^2 ;$$

$$\sigma^2_q = \frac{1}{GHSB(P-1)} \sum_{ijkmn} q^2_{ijkmn} .$$

We note that

$$GH\sigma^2_q = (G-1)(H-1)\,\Omega^2_{ghp} + G(H-1)\,\Omega^2_{hp} + (G-1)H\,\Omega^2_{gp} .$$

5. Case 1 - additivity of treatment factors with units, blocks and sources

We shall first give some results for the split plot design under simplifying assumptions. In this section we restrict ourself with the assumptions that the treatment combinations are additive with respect to the experimental units for the given scale of the observations, i.e. that

$$Y_{ijkmn} = X_{ijk} + W_{mn},$$

where X_{ijk} depends on the experimental unit while W_{mn} depends only on the treatment combination. This implies that

$$0 = (sg)_{im} = (sh)_{in} = (sgh)_{imn} = (bg)_{ijm} = (bh)_{ijn}$$

$$= (bgh)_{ijmn} = q_{ijkmn}, \text{ for all } i, j, k, m, n.$$

Under these conditions, the replication represented by d and r, and denoted by the subscripts f and e, is unnecessary. Thus in this section we take r = d = 1 and do not write the subscripts f and e. The sums of squares B', I'_{BH} and P' of Table 31 will not exist.

The statistical model then becomes

$$x_{i^*m^*n^*} = \mu + \sum_i a_i^{i^*} s_i + \sum_m \gamma_m^{m^*} g_m + \sum_n \delta_n^{n^*} h_n$$

$$+ \sum_{ij} a_i^{i^*} \beta_{i^*j}^{i^*m^*} b_{ij} + \sum_{mn} \gamma_m^{m^*} \delta_n^{n^*} (gh)_{mn}$$

$$+ \sum_{ijk} a_i^{i^*} \beta_{i^*j}^{i^*m^*} \rho_{i^*m^*k}^{i^*m^*n^*} p_{ijk} + e_{i^*m^*n^*}^* \quad .$$

This model could be written in the form

$$x_{i^*m^*n^*} = \mu + s_{i^*}^* + g_{m^*}^* + h_{n^*}^* + b_{i^*m^*}^* + (gh)_{m^*n^*}^* + e_{i^*m^*n^*}^* \quad ,$$

with the obvious correspondences. This latter form is reminiscent of the usual model employed in the analysis of the split plot design.

The expectations of the analysis of variance mean squares under the conditions of Case 1 are given in Table 32.

It will be seen from Table 32 that, under the conditions of Case 1, I^*_{SH} and I^*_{SGH} are estimates of the same quantity, namely $\sigma^2 + \sigma_p^2$. Thus we might "pool" these as follows:

$$\text{Let } M^* = \frac{(s-1)(h-1)\, I^*_{SH} + (s-1)(g-1)(h-1)\, I^*_{SGH}}{(s-1)g(h-1)}$$

$$= \frac{I^*_{SH} + (g-1)\, I^*_{SGH}}{g} \quad ;$$

then

$$E(M^*) = \sigma^2 + \sigma_p^2 \quad .$$

Table 32. Expected mean squares for split plot design under the conditions of Case 1

Mean Square	Expected Mean Square
S*	$\sigma^2 + \frac{(P-h)}{P}\sigma_p^2 + \frac{(B-g)}{B}h\sigma_b^2 + gh\sigma_s^2$
G*	$\sigma^2 + \frac{(P-h)}{P}\sigma_p^2 + h\sigma_b^2 + \frac{(H-h)}{H}s\sigma_{gh}^2 + sh\sigma_g^2$
I^*_{SG}	$\sigma^2 + \frac{(P-h)}{P}\sigma_p^2 + h\sigma_b^2$
H*	$\sigma^2 + \sigma_p^2 + \frac{(G-g)}{G}s\sigma_{gh}^2 + sg\sigma_h^2$
I^*_{GH}	$\sigma^2 + \sigma_p^2 + s\sigma_{gh}^2$
I^*_{SH}	$\sigma^2 + \sigma_p^2$
I^*_{SGH}	$\sigma^2 + \sigma_p^2$.

Unbiased estimates of the components of variation σ_g^2, σ_h^2 and σ_{gh}^2 can be obtained from the analysis of variance mean squares under the conditions of Case 1 whatever the relations of P, h, G, g, H, h, and B, g.

Thus

$$\hat{\sigma}_{gh}^2 = \frac{1}{s}(I^*_{GH} - M^*)$$

is an unbiased estimate of σ_{gh}^2;

$$\hat{\sigma}^2_h = \frac{1}{sg}\left[H^* - \frac{(G-g)}{G}\, s\hat{\sigma}^2_{gh} - M^*\right]$$

is an unbiased estimate of σ^2_h ;

$$\hat{\sigma}^2_g = \frac{1}{sh}\left[G^* - \frac{(H-h)}{H}\, s\hat{\sigma}^2_{gh} - I^*_{SG}\right]$$

is an unbiased estimate of σ^2_g. An unbiased estimate of σ^2_s is possible only when $B \gg g$.

It is of interest to note that the analysis of variance usually presented for the split plot design (e.g. in Cochran and Cox (1950), or Kempthorne (1952a)) is obtained from Case 1 here by taking:

$$G = g,$$
$$H = h,$$

and sources are taken as replicates. The sums of squares for I^*_{GH} and I^*_{SGH} are pooled, as would be natural if they have the same expected mean square. Whether these conditions are realistic we shall not discuss. They serve only to indicate how a correspondence can be made. It then turns out that with the usual terminology

$$\text{error (b)} = \sigma^2 + \sigma^2_p ,$$

$$\text{error (a)} = \sigma^2 + \sigma^2_p + h(\sigma^2_b - \frac{1}{P}\sigma^2_p) ,$$

so that the pure split plot error component is

$$\sigma^2 + \sigma^2_p$$

and the additional whole plot error component is

$$(\sigma^2_b - \frac{1}{P}\sigma^2_p) .$$

When the treatment factors are "fixed", there may be interest in estimates of contrasts such as $\sum_m c_m g_m$, where $\sum_m c_m = 0$. If we define (when G = g)

$$\hat{g}_m = x_{.m.} - x_{...}$$

$$\hat{h}_n = x_{..n} - x_{...}$$

$$(\hat{gh})_{mn} = x_{.mn} - x_{.m.} - x_{..n} + x_{...},$$

then $\hat{g}_m$, $\hat{h}_n$, $(\hat{gh})_{mn}$ are unbiased estimates of g_m, h_n, and $(gh)_{mn}$ respectively. Further an unbiased estimate of the treatment mean $Y_{...mn}$ is given by $x_{.mn}$.

If we only specify G = g, then $\hat{g}_m$ is still an unbiased estimate of g_m. Under the conditions of Case 1, the average variance of differences like $(\hat{g}_m - \hat{g}_{m'})$ is

$$\frac{2}{sh}\left[E(G^*) - sh\sigma_g^2\right] = 2\left[\frac{\sigma^2 + \sigma_P^2}{sh} + \frac{1}{s}(\sigma_b^2 - \frac{1}{P}\sigma_p^2) + \frac{(H-h)}{Hh}\sigma_{gh}^2\right].$$

An unbiased estimate of this variance is given by

$$\frac{2}{sh}\left[I^*_{SG} + \frac{(H-h)}{H}(I^*_{GH} - M^*)\right],$$

and this becomes

$$\frac{2}{sh} I^*_{SG}, \text{ when } H = h,$$

and

$$\frac{2}{sh}(I^*_{GH} - M^*), \text{ when } H >> h.$$

If also H = h, then the variance of $\sum_m c_m \hat{g}_m$, $\Sigma c_m = 0$, is

$$\left(\sum_m c_m^2\right)\left[\frac{\sigma^2+\sigma_p^2}{sh}+\frac{1}{s}\left(\sigma_b^2-\frac{1}{P}\sigma_p^2\right)\right],$$

which is estimated unbiasedly by

$$\frac{\left(\sum_m c_m^2\right)}{sh} I^*_{SG}.$$

If we specify $H = h$, then $\sum_n c_n \hat{h}_n$ is an unbiased estimate of $\sum_n h_n$. The average variance of differences such as $(\hat{h}_n - \hat{h}_{n'})$ is

$$\frac{2}{sg}\left(\sigma^2+\sigma_p^2+\frac{(G-g)}{G}s\sigma_{gh}^2\right),$$

and this is estimated unbiasedly by

$$\frac{2}{sg}\left[M^*+\frac{(G-g)}{G}\left(I^*_{G^*}-M^*\right)\right],$$

which becomes

$$\frac{2}{gh} M^*, \text{ when } G = g,$$

and

$$\frac{2}{sg} I^*_{GH}, \text{ when } G \gg g.$$

If also $G = g$, then the variance of $\sum_n c_n \hat{h}_n$, $\sum_n c_n = 0$, is

$$\frac{\left(\sum_n c_n^2\right)}{sg}\left(\sigma^2+\sigma_p^2\right),$$

which is estimated unbiasedly by

$$\frac{\left(\sum_n c_n^2\right)}{sg} M^*.$$

6. Case 2 - non-additivity of sources and treatment factors

The results of this section are valid under somewhat more general conditions than those of Case 1 in that we drop the restriction of additivity of sources with treatment factors. However we retain the assumptions of additivity of treatment combinations with units and with blocks, i.e. that

$$0 = (bg)_{ijm} = (bh)_{ijn} = (bgh)_{ijmn} = q_{ijmn}, \text{ for all } i, j, k, m, n.$$

Under these conditions the replication of levels of $\mathcal{A}$ within blocks is unnecessary but the replication of levels of $\mathcal{G}$ within sources would be useful. Hence we will give results for the case $r = 1$, $d > 1$.

Thus the appropriate statistical model for Case 2 is obtained from the general one developed above by dropping the subscript e everywhere and eliminating the components involving $(bg)_{ijm}$, $(bh)_{ijn}$, $(bgh)_{ijmn}$, and q_{ijkmn}. The line labelled $\mathcal{P}$ in Table 31 will not exist when $r = 1$.

The expectations of the analysis of variance mean squares under the conditions of Case 2 are given in Table 33. By inspection of Table 33, it may be seen that, with $d > 1$ and under the conditions of Case 2, the components of variation σ^2_{sgh}, σ^2_{sh}, σ^2_{gh}, σ^2_{sg}, σ^2_{h}, and σ^2_{g} are estimable from the analysis of variance mean squares whatever the relations of population sizes to sample sizes.

The role of the within source replication of levels of $\mathcal{G}$ is evident from Table 34 which lists "proper error terms". The meaning and use of these error terms is exemplified for factor $\mathcal{G}$ as follows; The error term V_g for the effects of factor $\mathcal{G}$ is a linear combination of

Table 33. Expected mean squares for split plot design under the conditions of Case 2.

Mean Square	Expected Mean Square
S*	$\sigma^2 + \frac{(P-h)}{P}\sigma^2_p + \frac{(B-gd)}{B}h\sigma^2_b + \frac{(G-g)}{G}\frac{(H-h)}{H}d\sigma^2_{sgh} + \frac{(G-g)}{G}dh\sigma^2_{sg} + \frac{(H-h)}{H}gd\sigma^2_{sh} + gdh\sigma^2_s$
G*	$\sigma^2 + \frac{(P-h)}{P}\sigma^2_p + h\sigma^2_b + \frac{(S-s)}{S}\frac{(H-h)}{H}d\sigma^2_{sgh} + \frac{(S-s)}{S}dh\sigma^2_{sg} + \frac{(H-h)}{H}sd\sigma^2_{gh} + sdh\sigma^2_g$
H*	$\sigma^2 + \sigma^2_p + \frac{(S-s)}{S}\frac{(G-g)}{G}d\sigma^2_{sgh} + \frac{(S-s)}{S}gd\sigma^2_{sh} + \frac{(G-g)}{G}sd\sigma^2_{gh} + sgd\sigma^2_h$
I^*_{SG}	$\sigma^2 + \frac{(P-h)}{P}\sigma^2_p + h\sigma^2_b + \frac{(H-h)}{H}d\sigma^2_{sgh} + dh\sigma^2_{sg}$
I^*_{GH}	$\sigma^2 + \sigma^2_p + \frac{(S-s)}{S}d\sigma^2_{sgh} + sd\sigma^2_{gh}$
I^*_{SH}	$\sigma^2 + \sigma^2_p + \frac{(G-g)}{G}d\sigma^2_{sgh} + gd\sigma^2_{sh}$
I^*_{SGH}	$\sigma^2 + \sigma^2_p + d\sigma^2_{sgh}$
B*	$\sigma^2 + \frac{(P-h)}{P}\sigma^2_p + h\sigma^2_b$
I^*_{BH}	$\sigma^2 + \sigma^2_p$

the analysis of variance mean squares if it exists, such that (i) an unbiased estimate of σ_g^2 is obtainable as a multiple of $(G^* - V_g)$, (ii) a meaningful "F-type" statistic is given by G^*/V_g, and (iii) if g is "fixed" then an unbiased estimate of the average variance of unbiased estimates, $(\hat{g}_m - \hat{g}_{m'})$, of differences $(g_m - g_{m'})$ is given by a multiple of V_g.

Table 34. Proper error terms. Case 2

$$V_{sgh} = I^*_{BH}$$

$$V_{sh} = \frac{g}{G} I^*_{BH} + \frac{(G-g)}{G} I^*_{SGH}$$

$$V_{gh} = \frac{s}{S} I^*_{BH} + \frac{(S-s)}{S} I^*_{SGH}$$

$$V_{sg} = B^* - \frac{(H-h)}{H} I^*_{BH} + \frac{(H-h)}{H} I^*_{SGH}$$

$$V_h = \frac{sg}{SG} I^*_{BH} - \frac{(S-s)}{S} \frac{(G-g)}{G} I^*_{SGH} + \frac{(G-g)}{G} I^*_{GH} + \frac{(S-s)}{S} I^*_{SH}$$

$$V_g = \frac{s}{S} B^* - \frac{s}{S} \frac{(H-h)}{H} I^*_{BH} - \frac{(S-s)}{S} \frac{(H-h)}{H} I^*_{SGH} + \frac{(S-s)}{S} I^*_{SG} + \frac{(H-h)}{H} I^*_{Gh}$$

For general relations of B, g, P, h, and d, a proper error term for sources does not exist. If, however, $B \gg gd$, then

$$V'_s = \frac{g}{G} B^* - \frac{g}{G} \frac{(H-h)}{H} I^*_{BH} - \frac{(G-g)}{G} \frac{(H-h)}{H} I^*_{SGH} + \frac{(G-g)}{G} I^*_{SG} + \frac{(H-h)}{H} I^*_{GH} .$$

If $P \gg h$, then, $V''_s = g(\frac{1}{G} - \frac{d}{B})B^* + g(\frac{d}{B} - \frac{1}{G} + \frac{h}{GH})I^*_{BH}$

$$- \frac{(G-g)}{G}\frac{(H-h)}{H} I^*_{SGH} + \frac{(G-g)}{G} I^*_{SG} + \frac{(H-h)}{H} I^*_{SH} \quad .$$

If the sources used in the experiment can be regarded as a random selection from a much larger population of possible sources, then under the conditions of Case 2 proper error terms for $\mathcal{G}$, $\mathcal{H}$ and $\mathcal{G} \times \mathcal{H}$ are obtainable without using the mean squares B^* or I^*_{BH}. Of course, even with $S \gg s$, treatment replication within blocks and within sources would still be necessary if the magnitude of interactions of sources with treatments were to be evaluated.

In many situations the different sources may exhibit considerable heterogeneity, since one of the objectives of the design is to eliminate variation due to source differences. Under such conditions the additivity or non-additivity of sources with treatments would be heavily dependent on the scale of observation. Thus the assumptions of Case 1 would often be too restrictive unless considerable prior information were available. The conditions of Case 2 may be more realistic in general if units within sources may be considered fairly homogeneous. However, we emphasize again that, in general, the magnitude and importance of interactions of treatments with experimental material, as between treatment factors, will depend on the scale of observation.

7. Case 3 - general conditions

We consider now some aspects of the analysis of the split plot design under the general conditions originally described, i.e. with no simplifying assumptions regarding any interaction terms. The

appropriate statistical model is the general one given above. We note that from some points of view it is desirable and meaningful to decompose the component

$$\sum_{ijkmn} a_i^{i*} \delta_n^{n*} \gamma_m^{m*} \beta_{i*j}^{i*m*f} \rho_{i*m*fk}^{i*m*fn*e} q_{ijkmn}$$

into three "orthogonal" components reflecting interactions of units within blocks within sources (i) with levels of $\mathcal{G}$ averaged over levels of $\mathcal{H}$, (ii) with levels of $\mathcal{H}$ averaged over levels of $\mathcal{G}$, and (iii) with interactions of $\mathcal{G}$ and $\mathcal{H}$. The decomposition is implicit in the analysis of variance, and would follow from the identity

$$q_{ijkmn} = q_{ijkm.} + q_{ijk.n} + (q_{ijkmn} - q_{ijkm.} - q_{ijk.n}) .$$

Within block replication of levels of $\mathcal{H}$ and within source replication of levels of $\mathcal{G}$ would be useful, so we give results for $d \geq 1$, $r \geq 1$.

The results in Table 35 on expected mean squares are presented in a form which indicates what are the estimable quantities in the analysis of variance under general conditions. The form also portrays a general pattern which makes extension of the results quite simple.* It is noteworthy that the Σ quantities in Table 35 depend solely on the population and not on the sample. The following notational definitions are used in Table 35:

---

*All the results on expected mean squares given in this thesis can be expressed in a form entirely analogous to that of Table 35. We have not done so everywhere in order not to depart too radically from usual notation and practice.

Table 35. Expected mean squares for split plot design under general conditions. Case 3

Mean Square	Expected Mean Square
P*	$\Sigma_o = \sigma^2 + \Sigma_p + \Sigma_{gp} + \Sigma_{hp} + \Sigma_{ghp}$
I^*_{BH}	$\Sigma_o + r\Sigma_{ghb} + r\Sigma_{hb}$
B*	$\Sigma_o + r\Sigma_{ghb} + r\Sigma_{hb} + hr\Sigma_{gb} + hr\Sigma_b$
I^*_{SGH}	$\Sigma_o + r\Sigma_{ghb} + r\Sigma_{hb} + dr\Sigma_{sgh}$
I^*_{GH}	$\Sigma_o + r\Sigma_{ghb} + r\Sigma_{hb} + dr\Sigma_{sgh} + sdr\Sigma_{gh}$
I^*_{SH}	$\Sigma_o + r\Sigma_{ghb} + r\Sigma_{hb} + dr\Sigma_{sgh} + gdr\Sigma_{sh}$
I^*_{SG}	$\Sigma_o + r\Sigma_{ghb} + r\Sigma_{hb} + hr\Sigma_{gb} + hr\Sigma_b + dr\Sigma_{sgh} + dhr\Sigma_{sg}$
H*	$\Sigma_o + r\Sigma_{ghb} + r\Sigma_{hb} + dr\Sigma_{sgh} + gdr\Sigma_{sh} + sdr\Sigma_{gh} + sgdr\Sigma_h$
G*	$\Sigma_o + r\Sigma_{ghb} + r\Sigma_{hb} + hr\Sigma_{gb} + hr\Sigma_b + dr\Sigma_{sgh} + sdr\Sigma_{gh} + dhr\Sigma_{sg} + sdhr\Sigma_g$
S*	$\Sigma_o + r\Sigma_{ghb} + r\Sigma_{hb} + hr\Sigma_{gb} + hr\Sigma_b + dr\Sigma_{sgh} + gdr\Sigma_{sh} + dhr\Sigma_{sg} + gdhr\Sigma_s$

$$\Sigma_{ghp} = Q^2_{ghp}$$

$$\Sigma_{hp} = (Q^2_{hp} - \frac{1}{G} Q^2_{ghp})$$

$$\Sigma_{gp} = (Q^2_{gp} - \frac{1}{H} Q^2_{ghp})$$

$$\Sigma_{p} = (\sigma^2_p - \frac{1}{H} Q^2_{hp} - \frac{1}{G} Q^2_{gp} + \frac{1}{GH} Q^2_{ghp})$$

$$\Sigma_{0} = (\Sigma_{ghp} + \Sigma_{hp} + \Sigma_{gp} + \Sigma_{p} + \sigma^2) = (\sigma^2 + \sigma^2_p + \sigma^2_q)$$

$$\Sigma_{ghb} = (Q^2_{ghb} - \frac{1}{P} Q^2_{ghp})$$

$$\Sigma_{hb} = (Q^2_{hb} - \frac{1}{G} Q^2_{ghb} - \frac{1}{P} Q^2_{hp} + \frac{1}{GP} Q^2_{ghp})$$

$$\Sigma_{gb} = (Q^2_{gb} - \frac{1}{H} Q^2_{ghb} - \frac{1}{P} Q^2_{gp} + \frac{1}{HP} Q^2_{ghp})$$

$$\Sigma_{b} = (\sigma^2_b - \frac{1}{H} Q^2_{hb} - \frac{1}{G} Q^2_{gb} - \frac{1}{P} \sigma^2_p + \frac{1}{HG} Q^2_{ghb} + \frac{1}{HP} Q^2_{hp} + \frac{1}{GP} Q^2_{gp} - \frac{1}{HGP} Q^2_{ghp})$$

$$\Sigma_{sgh} = (\sigma^2_{sgh} - \frac{1}{B} Q^2_{ghb})$$

$$\Sigma_{sh} = (\sigma^2_{sh} - \frac{1}{G} \sigma^2_{sgh} - \frac{1}{B} Q^2_{hb} + \frac{1}{GB} Q^2_{ghb})$$

$$\Sigma_{sg} = (\sigma^2_{sg} - \frac{1}{H} \sigma^2_{sgh} - \frac{1}{B} Q^2_{gb} + \frac{1}{HB} Q^2_{ghb})$$

$$\Sigma_{gh} = (\sigma^2_{gh} - \frac{1}{S} \sigma^2_{sgh})$$

$$\Sigma_{s} = \sigma^2_s - \frac{1}{G} \sigma^2_{sg} - \frac{1}{H} \sigma^2_{sh} - \frac{1}{B} \sigma^2_b + \frac{1}{GH} \sigma^2_{sgh} + \frac{1}{GB} Q^2_{gb} + \frac{1}{HB} Q^2_{hb} - \frac{1}{GHB} Q^2_{ghb}$$

$$\Sigma_{g} = \sigma^2_g - \frac{1}{S} \sigma^2_{sg} - \frac{1}{H} \sigma^2_{gh} + \frac{1}{SH} \sigma^2_{sgh}$$

$$\Sigma_{h} = \sigma^2_h - \frac{1}{S} \sigma^2_{sh} - \frac{1}{G} \sigma^2_{gh} + \frac{1}{SG} \sigma^2_{sgh} \ .$$

From their definition we see that as S, B, P, G, and H all become large each of the Σ's (except Σ_0) approaches its corresponding σ^2 or Q^2 quantity. Thus the form of Table 35 would correspond to a "usual" table of expected mean squares with "everything random".

One might be tempted, from the structure of Table 35, to infer that it is in fact the Σ's which one is concerned with estimating. This will not in general be so. For example: by definition

$$s_i = (Y_{i\ldots\ldots} - Y_{\ldots\ldots}),$$

which is the difference between the response on source i, averaged over all treatments, and the average response over all sources. An unbiased estimate of s_i (when $S = s$) is provided by

$$\hat{s}_i = x_{i\ldots\ldots} - x_{\ldots\ldots}.$$

Now a reasonable measure of the variation or dispersion of the population of s_i values is the average value of squares of differences such as $(s_i - s_{i'})$, $i \neq i'$. This is

$$\frac{1}{S(S-1)} \sum_{i \neq i'} (s_i - s_{i'})^2 = 2\sigma_s^2 .$$

Hence as a measure of the variation among sources, with respect to responses averaged over treatments, it seems reasonable indeed to use the component of variation σ_s^2.

As something of an aside we note that, by definition of variance of a random variable, the variance of

$$\hat{s}^*_{i*} = \sum_i a_i^{i*} s_i, \text{ when } S > s,$$

is

$$E(s^*_{i*})^2 = E\left(\sum_i a_i^{i*} s_i^2\right) = \frac{1}{S}\sum_i s_i^2 = \frac{S-1}{S}\sigma_s^2 \quad .$$

Thus, the variance of the component s^*_{i*} of the statistical model is not the same as the defined component of variation of the population $\{s_i\}$, though the two are equal in the limit as S gets very large.

We turn our attention now to the estimation of components of variation and of the variance of estimates under general conditions. First we check on the properties of some Case 1 and Case 2 estimators when in fact Case 3 conditions obtain.

Under Case 1 conditions an unbiased estimate of σ^2_{gh} would be given by

$$\frac{1}{sdr}\left(I^*_{GH} - \frac{I^*_{SH} + (g-1)I^*_{SGH}}{g}\right) .$$

Under general conditions, we see from Table 35 that this would have expectation

$$\Sigma_{gh} - \frac{(g-1)}{s}\Sigma_{sh}$$

$$= \sigma^2_{gh} - \frac{(g-1)}{s}\sigma^2_{sh} - \left(\frac{1}{S} - \frac{(g-1)}{sG}\right)\sigma^2_{sgh} + \frac{(g-1)}{Bs}\left(\Omega^2_{hb} - \frac{1}{G}\Omega^2_{ghb}\right).$$

Evidently the bias in this estimate of σ^2_{gh} may in general be serious. For example, with S and B large, and G = g, the bias would be

$$-\frac{(g-1)}{s}\left(\sigma^2_{sh} - \frac{1}{g}\sigma^2_{sgh}\right) .$$

From the fact that this bias will in general be negative, one might infer that, in using

$$\frac{I^*_{SH} + (g-1)I^*_{SGH}}{g}$$

as an error term in a test of significance of $\mathcal{G} \times \mathcal{H}$ interactions, there would be a tendency to underestimate significance.

Under the conditions of Case 2 a proper error term for $\mathcal{G} \times \mathcal{H}$ interactions was V_{gh}, defined in Table 34. Under general conditions

$$E(V_{gh}) = \Sigma_0 + \frac{(S-s)}{S} dr \ \Sigma_{sgh}$$

and

$$\frac{E(I^*_{GH} - V_{gh})}{srd} = \Sigma_{gh} + \frac{1}{S} \Sigma_{sgh}$$

$$= \sigma^2_{gh} - \frac{1}{SB} Q^2_{ghb} \ .$$

This bias will certainly be unimportant when S or B are large. (Note that relative values of S and s or B and gd do not enter.) When SB is not large the bias may still be unimportant if blocks within sources are reasonably homogeneous, for then the interactions of blocks with $\mathcal{G} \times \mathcal{H}$ will tend to be small for most scales of observation. Since the bias is always negative one would infer that the use of V_{gh} as an error term for $\mathcal{G} \times \mathcal{H}$ would tend to underestimate significance.

It is of interest to enquire whether, having all the mean squares of Table 31 and the information of Table 35, we can find an estimate of σ^2_{gh} with a smaller bias than the Case 2 estimate. To facilitate the use of Table 35 for this and other purposes we give in Table 36 the various σ^2 and Q^2 quantities as explicit functions of the Σ's. In using using Table 35 it is Table 36, rather than the definitions of the Σ's, which will be most useful. The pattern of the relation of the σ^2's and

Table 36. Relations between σ^2's, Q^2's and Σ's.

$$\sigma_s^2 = \Sigma_s + \frac{1}{G}\Sigma_{sg} + \frac{1}{H}\Sigma_{sh} + \frac{1}{GH}\Sigma_{sgh} + \frac{1}{B}\Sigma_b + \frac{1}{GB}\Sigma_{gb}$$
$$+ \frac{1}{HB}\Sigma_{hb} + \frac{1}{GHB}\Sigma_{ghb} + \frac{1}{BP}\Sigma_p + \frac{1}{GBP}\Sigma_{gp}$$
$$+ \frac{1}{HBP}\Sigma_{hp} + \frac{1}{GHBP}\Sigma_{ghp}$$

$$\sigma_g^2 = \Sigma_g + \frac{1}{S}\Sigma_{sg} + \frac{1}{H}\Sigma_{gh} + \frac{1}{SH}\Sigma_{sgh} + \frac{1}{SB}\Sigma_{gb} + \frac{1}{SBH}\Sigma_{ghb}$$
$$+ \frac{1}{SBP}\Sigma_{gp} + \frac{1}{SBPH}\Sigma_{ghp}$$

$$\sigma_h^2 = \Sigma_h + \frac{1}{S}\Sigma_{sh} + \frac{1}{G}\Sigma_{gh} + \frac{1}{SG}\Sigma_{sgh} + \frac{1}{SB}\Sigma_{hb}$$
$$+ \frac{1}{SBG}\Sigma_{ghb} + \frac{1}{SBP}\Sigma_{hp} + \frac{1}{SBPG}\Sigma_{ghp}$$

$$\sigma_{sg}^2 = \Sigma_{sg} + \frac{1}{H}\Sigma_{sgh} + \frac{1}{B}\Sigma_{gb} + \frac{1}{BH}\Sigma_{ghb} + \frac{1}{BP}\Sigma_{gp}$$
$$+ \frac{1}{BPH}\Sigma_{ghp}$$

$$\sigma_{sh}^2 = \Sigma_{sh} + \frac{1}{G}\Sigma_{sgh} + \frac{1}{B}\Sigma_{hb} + \frac{1}{BG}\Sigma_{ghb} + \frac{1}{BP}\Sigma_{hp}$$
$$+ \frac{1}{BPG}\Sigma_{ghp}$$

$$\sigma_{gh}^2 = \Sigma_{gh} + \frac{1}{S}\Sigma_{sgh} + \frac{1}{SB}\Sigma_{ghb} + \frac{1}{SBP}\Sigma_{ghp}$$

Table 36 (Continued)

$$\sigma^2_{sgh} = \Sigma_{sgh} + \frac{1}{B}\Sigma_{ghb} + \frac{1}{BP}\Sigma_{ghp}$$

$$\sigma^2_{b} = \Sigma_{b} + \frac{1}{G}\Sigma_{gb} + \frac{1}{H}\Sigma_{hb} + \frac{1}{GH}\Sigma_{ghb} + \frac{1}{P}\Sigma_{p} + \frac{1}{GP}\Sigma_{gp} + \frac{1}{HP}\Sigma_{hp} + \frac{1}{GHP}\Sigma_{ghp}$$

$$Q^2_{gb} = \Sigma_{gb} + \frac{1}{H}\Sigma_{ghb} + \frac{1}{P}\Sigma_{gp} + \frac{1}{PH}\Sigma_{ghp}$$

$$Q^2_{hb} = \Sigma_{hb} + \frac{1}{G}\Sigma_{ghb} + \frac{1}{P}\Sigma_{hp} + \frac{1}{PG}\Sigma_{ghp}$$

$$Q^2_{ghb} = \Sigma_{ghb} + \frac{1}{P}\Sigma_{ghp}$$

$$\sigma^2_{p} = \Sigma_{p} + \frac{1}{G}\Sigma_{gp} + \frac{1}{H}\Sigma_{hp} + \frac{1}{GHP}\Sigma_{ghp}$$

$$Q^2_{gp} = \Sigma_{gp} + \frac{1}{H}\Sigma_{ghp}$$

$$Q^2_{hp} = \Sigma_{hp} + \frac{1}{G}\Sigma_{ghp}$$

$$Q^2_{ghp} = \Sigma_{ghp}$$

Q^2's to the Σ's will be apparent.

Turning back now to the estimation of σ^2_{gh}, we have from Table 36

$$\sigma^2_{gh} = \Sigma_{gh} + \frac{1}{S}\Sigma_{sgh} + \frac{1}{SB}\Sigma_{ghb} + \frac{1}{SBP}\Sigma_{ghp} .$$

Hence we try to find from Table 35 unbiased estimates of the appropriate Σ's. An unbiased estimate of Σ_{gh} is

$$\hat{\Sigma}_{gh} = \frac{1}{sdr}(I^*_{GH} - I^*_{SGH}) ,$$

and an unbiased estimate of Σ_{sgh} is

$$\hat{\Sigma}_{sgh} = \frac{1}{dr}(I^*_{SGH} - I^*_{BH}) .$$

The quantity $(\Sigma_{ghb} + \frac{1}{P}\Sigma_{ghp})$ is not estimable. If we use

$$\hat{\Sigma}_{gh} + \frac{1}{S}\hat{\Sigma}_{sgh}$$

as an estimate of σ^2_{gh}, this is precisely the Case 2 estimate using V_{gh}. If we take

$$\frac{1}{r}(I^*_{BH} - P^*)$$

as an estimate of $(\Sigma_{ghb} + \frac{1}{P}\Sigma_{ghp})$, and investigate

$$\tilde{\sigma}^2_{gh} = \hat{\Sigma}_{gh} + \frac{1}{S}\hat{\Sigma}_{sgh} + \frac{1}{SBr}(I^*_{BH} - P^*) ,$$

we find that

$$E(\tilde{\sigma}^2_{gh}) = \Sigma_{gh} + \frac{1}{S}\Sigma_{sgh} + \frac{1}{SB}(\Sigma_{ghb} + \Sigma_{hb})$$

$$= \sigma^2_{gh} + \frac{1}{SB}(Q^2_{hb} - \frac{1}{G}Q^2_{ghb} - \frac{1}{P}Q^2_{hp} - \frac{1}{P}Q^2_{ghp} + \frac{1}{GP}Q^2_{ghp}) .$$

The bias of $\tilde{\sigma}^2_{gh}$ will in general be positive and would in many cases exceed in magnitude the bias from the Case 2 estimate, since Q^2_{hb} represents two-way interactions, while Q^2_{ghb} reflects three-way interactions. The use of the error term corresponding to $\tilde{\sigma}^2_{gh}$ in a significance test would tend to overestimate significance.

We consider now the estimation of differences of (main) effects of levels of $\mathcal{G}$ i.e. quantities such as $(g_m - g_{m'})$, when $G = g$. If we define

$$\hat{g}_m = x_{.m\ldots} - x_{\ldots\ldots}$$

then it is easy to check from the general statistical model that $(\hat{g}_m - \hat{g}_{m'})$ is an unbiased estimate of $(g_m - g_{m'})$ under the general conditions of Case 3. The average variance of estimates of such differences is

$$\frac{1}{G(G-1)} \sum_{m \neq m'} E(x_{.m\ldots} - x_{.m'\ldots} - g_m + g_{m'})^2$$

$$= \frac{E(G^*)}{sdhr} - \sigma^2_g \quad .$$

To estimate this average variance we require a "proper error term" for $\mathcal{G}$.

Under Case 1 we would use

$$\hat{\sigma}^2_g = \frac{1}{sdhr} \left[G^* - \frac{(H-h)}{H} \left(\frac{I^*_{GH} - I^*_{SH} + (g-1)\, I^*_{SGH}}{g} \right) - I^*_{SG} \right] ;$$

and under Case 2 we would use

$$\hat{\hat{\sigma}}^2_g = \frac{1}{sdhr} \left[G^* - V_g \right].$$

Now, under general conditions

$$E(\hat{\sigma}^2_g) = \Sigma_g + \frac{1}{H}\Sigma_{gh} + \frac{(H-h)}{H_s}\Sigma_{sh}$$

$$= \sigma^2_g - \frac{1}{S}\sigma^2_{sg} + \frac{(H-h)}{H_s}(\sigma^2_{sh} - \frac{1}{G}\sigma^2_{sgh} - \frac{1}{B}\Omega^2_{hb} + \frac{1}{GB}\Omega^2_{ghb}) ;$$

and,

$$E(\hat{\hat{\sigma}}^2_g) = \sigma^2_g - \frac{1}{SB}\Omega^2_{gb} .$$

Since $\hat{\sigma}^2_g$ will usually be negatively biased and $\hat{\hat{\sigma}}^2_g$ will always be negatively biased the use of either the Case 1 or the Case 2 error term will result in overestimation of the average variance of the estimates $(\hat{g}_m - \hat{g}_{m'})$. Of course the Case 2 estimate of this average variance, namely

$$\frac{2}{sdhr} V_g ,$$

will usually be the more desirable of the two in that its bias will be substantially less.

It is easy to check that using V_g corresponds to taking, as an estimate of σ^2_g,

$$\tilde{\sigma}^2_g = \hat{\Sigma}_g + \frac{1}{S}\hat{\Sigma}_{sg} + \frac{1}{H}\hat{\Sigma}_{gh} + \frac{1}{SH}\hat{\Sigma}_{sgh} ,$$

where $\hat{\Sigma}_g = \frac{1}{sdhr}(G^* - B^* - dr\,\hat{\Sigma}_{sgh} - sdr\hat{\Sigma}_{gh} - dhr\hat{\Sigma}_{sg})$,

$$\hat{\Sigma}_{sg} = \frac{1}{dhr}(I^*_{SG} - B^* - dr\hat{\Sigma}_{sgh})$$

$$\hat{\Sigma}_{gh} = \frac{1}{sdr}(I^*_{GH} - I^*_{SGH})$$

$$\hat{\Sigma}_{sgh} = \frac{1}{dr}(I^*_{SGH} - I^*_{BH}) .$$

The $\hat{\Sigma}$'s are unbiased estimates of the corresponding Σ's. Since the remaining Σ's in the expression for σ_g^2 in Table 36, namely Σ_{gb}, Σ_{ghb}, Σ_{gp}, Σ_{ghp}, are not estimable the error term V_g for $\mathcal{G}$ is about as good as we can do.

We consider now the specification of an error term for $\mathcal{H}$. Under Case 1 conditions we would use

$$\hat{\sigma}_h^2 = \frac{1}{sgdr}\left[H^* - \frac{(G-g)}{G} I^*_{GH} - \frac{1}{G} I^*_{SH} - \frac{(g-1)}{G} I^*_{SGH}\right],$$

while under Case 2 conditions we would use

$$\hat{\hat{\sigma}}_h^2 = \frac{1}{sgdr}\left[H^* - V_h\right],$$

as estimates of σ_h^2.

Now, under general conditions,

$$E(\hat{\hat{\sigma}}_h^2) = \Sigma_h + \frac{1}{G}\Sigma_{gh} + \frac{1}{s}\left(1 - \frac{1}{G}\right)\Sigma_{sh}$$

$$= \sigma_h^2 + \left(\frac{1}{s} - \frac{1}{sG} - \frac{1}{S}\right)\sigma_{sh}^2 - \frac{1}{s}\left(1 - \frac{1}{G}\right)\left(\frac{1}{G}\sigma_{sgh}^2 + \frac{1}{B}\Omega_{hb}^2 - \frac{1}{GB}\Omega_{ghb}^2\right).$$

The bias involved will usually be positive and may be important.

For the estimate based on V_h we find

$$E(\hat{\hat{\sigma}}_h^2) = \Sigma_h + \frac{1}{S}\Sigma_{sh} + \frac{1}{G}\Sigma_{gh} + \frac{1}{SG}\Sigma_{sgh}$$

$$= \sigma_h^2 - \frac{1}{SB}\Omega_{hb}^2 .$$

The bias of $\hat{\hat{\sigma}}_h^2$ is negative and hence the use of V_h as an error term will tend to underestimate significance; and will tend to overestimate the average variance of estimates, $(\hat{h}_n - \hat{h}_{n'})$, of the differences, $(h_n - h_{n'})$, when $H = h$, where $\hat{h}_n = x_{\ldots n.} - x_{\ldots n'.}$ However, if S or B are large

the bias becomes unimportant, even if the interactions of levels of $\mathcal{H}$ with blocks within sources are important.

We might consider, as an estimate of σ_h^2, $\tilde{\sigma}_h^2 = \hat{\Sigma}_h + \frac{1}{S}\hat{\Sigma}_{sh} + \frac{1}{G}\hat{\Sigma}_{gh} + \frac{1}{SG}\hat{\Sigma}_{sgh} + \frac{1}{SBr}(I^*_{BH} - P^*)$, where each $\hat{\Sigma}$ is an unbiased estimate of the corresponding Σ and, it turns out that the sum of the first four terms is V_h. We find that

$$E(\tilde{\sigma}_h^2) = \sigma_h^2 - \frac{1}{SB}\Omega_{hb}^2 + \frac{1}{SB}(\Sigma_{ghb} + \Sigma_{hb})$$

$$= \sigma_h^2 + \frac{1}{SB}\left[\left(1 - \frac{1}{G}\right)\Omega_{ghb}^2 - \frac{1}{P}\Omega_{hp}^2 - \frac{1}{P}\left(1 - \frac{1}{G}\right)\Omega_{ghp}^2\right].$$

The bias of $\tilde{\sigma}_h^2$ will usually be positive, and will ordinarily be smaller in magnitude than the bias of $\hat{\hat{\sigma}}_h^2$.

It is to be noted that V_h (and hence $\hat{\hat{\sigma}}_h^2$) does not depend on P*, while $\tilde{\sigma}_h^2$ (and the corresponding error term) does depend on P*. In most circumstances the decrease in bias in using $\tilde{\sigma}_h^2$ instead of $\hat{\hat{\sigma}}_h^2$ as an estimate of σ_h^2 would not justify having replication of levels of $\mathcal{H}$ within blocks. Unless there is interest in evaluating possible interactions of treatment combinations with blocks within sources there would not be much lost in taking r = 1. However it does appear that, even if the evaluation of interactions of treatments with sources is not of interest, having replication of levels of $\mathcal{G}$ within sources would be quite advantageous.

We close this section with an illustration of the use of Table 35 and its mate Table 36, in finding an estimate for σ_s^2. It is clear that unbiased estimates of Σ_s, Σ_{sg}, Σ_{sh} and Σ_{sgh} are

$$\hat{\Sigma}_s = \frac{1}{gdhr}(S^* - I^*_{SG} - gdr\,\hat{\Sigma}_{sh}) ,$$

$$\hat{\Sigma}_{sg} = \frac{1}{dhr}(I^*_{SG} - B^* - dr\,\hat{\Sigma}_{sgh}) ,$$

$$\hat{\Sigma}_{sh} = \frac{1}{gdr}(I^*_{SH} - I^*_{SGH}) ,$$

$$\hat{\Sigma}_{sgh} = \frac{1}{dr}(I^*_{SGH} - I^*_{BH}) ,$$

respectively. Further

$$\hat{\hat{\Sigma}}_b = \frac{1}{hr}(B^* - I^*_{BH})$$

is an unbiased estimate of $(\Sigma_b + \Sigma_{gb})$; while

$$\hat{\hat{\Sigma}}_{hb} = \frac{1}{r}(I^*_{BH} - P^*)$$

is an unbiased estimate of $(\Sigma_{hb} + \Sigma_{ghb})$. Finally, if σ^2 is small compared with σ^2_p, Ω^2_{hp}, etc., then we might use P^* as an approximate estimate of Σ_p.

Thus let

$$\tilde{\sigma}^2_s = \hat{\Sigma}_s + \frac{1}{G}\hat{\Sigma}_{sg} + \frac{1}{H}\hat{\Sigma}_{sh} + \frac{1}{GH}\hat{\Sigma}_{sgh} + \frac{1}{B}\hat{\hat{\Sigma}}_b + \frac{1}{HB}\hat{\hat{\Sigma}}_{hb} + \frac{1}{BP}P^* ,$$

ignoring the remaining Σ components of σ^2_s in Table 36 since they are not estimable. We then find that

$$\begin{aligned} E(\tilde{\sigma}^2_s) &= \Sigma_s + \frac{1}{G}\Sigma_{sg} + \frac{1}{H}\Sigma_{sh} + \frac{1}{GH}\Sigma_{sgh} + \frac{1}{B}\Sigma_b + \frac{1}{B}\Sigma_{gb} \\ &\quad + \frac{1}{HB}\Sigma_{hb} + \frac{1}{HB}\Sigma_{ghb} + \frac{1}{BP}\Sigma_p + \frac{1}{BP}(\Sigma_0 - \Sigma_p) \\ &= \sigma^2_s + \frac{1}{B}(1 - \frac{1}{G})\Sigma_{gb} + \frac{1}{HB}(1 - \frac{1}{G})\Sigma_{ghb} + \frac{1}{BP}(1 - \frac{1}{G})\Sigma_{gp} \\ &\quad + \frac{1}{BP}(1 - \frac{1}{H})\Sigma_{hp} + \frac{1}{BP}(1 - \frac{1}{GH})\Sigma_{ghp} + \frac{1}{BP}\sigma^2 \end{aligned}$$

$$= \sigma_s^2 + \frac{1}{B}(1 - \frac{1}{G})Q_{gb}^2 + \frac{1}{BP}\left[\sigma^2 + (1 - \frac{1}{H})Q_{hp}^2 + (1 - \frac{1}{G} - \frac{1}{H} + \frac{1}{GH})Q_{ghp}^2\right].$$

8. Case 4 - generalization of split plot results

The natural extension of the split plot design is the so-called split split plot design. In this latter design the experimental units are structured according to a four-fold hierarchy, say units (P) within blocks (B) within sources (S) within location (L), while the treatments are classified according to three factors $\mathcal{F}$, $\mathcal{G}$ and $\mathcal{H}$. Levels of $\mathcal{F}$ are applied to sources, levels of $\mathcal{G}$ to blocks, and levels of $\mathcal{H}$ to units.

Let there be: L location,

S sources per location,

B blocks per source,

P units per block,

F levels of $\mathcal{F}$,

G levels of $\mathcal{G}$,

H levels of $\mathcal{H}$.

The experimental procedure and design is the obvious extension of that for the simple split plot.

Our conceptual true yield is now $Y_{uijkvmn}$ where

u denotes the location,

i denotes source within location,

j denotes blocks within source,

k denotes unit within block,

v denotes level of $\mathcal{F}$,

m denotes level of $\mathcal{G}$,

n denotes level of $\mathcal{H}$.

The definition of population parameters follows the same pattern as for the simple split plot, for example

$$\mu = Y_{.....}$$

$$l_u = Y_{u.....} - \mu$$

$$s_{ui} = Y_{ui.....} - Y_{u.....}$$

$$f_v = Y_{....v..} - \mu$$

$$g_m = Y_{.....m.} - \mu$$

etc.

The population model we take as

$$\begin{aligned} y_{uijkvmn} = \mu &+ l_u + f_v + g_m + h_n + (lf)_{uv} + (lg)_{um} + (lh)_{un} \\ &+ (fg)_{vm} + (fh)_{vn} + (gh)_{mn} + (lfg)_{uvm} \\ &+ (lgh)_{umn} + (fgh)_{vmn} + (lfgh)_{uvmn} + s_{ui} \\ &+ \cdots\cdots + (sfgh)_{uivmn} + b_{uij} + \cdots\cdots + (bfgh)_{uijvmn} \\ &+ p_{uijk} + (pf)_{uijkv} + \cdots\cdots + (pfgh)_{uijkvmn} + \epsilon_{uijkvmn} . \end{aligned}$$

The definition of components of variation proceeds as usual, for example

$$\sigma_l^2 = \frac{1}{(L-1)} \sum_u l_u^2 ,$$

$$\sigma_f^2 = \frac{1}{(F-1)} \sum_v f_v^2 ,$$

$$\sigma_{lg}^2 = \frac{1}{(L-1)(G-1)} \sum_{um} (lg)_{um}^2 ,$$

$$\sigma_s^2 = \frac{1}{L(S-1)} \sum_{ui} s_{ui}^2 ,$$

$$\sigma_b^2 = \frac{1}{LS(B-1)} \sum_{uij} b_{uij}^2 ,$$

$$\sigma_p^2 = \frac{1}{LSB(P-1)} \sum_{uijk} p_{uijk}^2 ,$$

$$\sigma_{pf}^2 = \frac{1}{LSB(P-1)(F-1)} \sum_{uijkv} (pf)_{uijkv}^2 ,$$

etc.

We suppose that in our experiment we select at random

f levels of $\mathcal{F}$, from F,

g levels of $\mathcal{G}$, from G,

h levels of $\mathcal{H}$, from H,

l locations from L,

f sources from each selected location,

g blocks from each selected source,

h units from each selected block.

(For simplicity in writing we have omitted the generalization to possible replication of, say, levels of $\mathcal{F}$ within locations.)

The analysis of variance would proceed by dividing the (lfgh - 1) degrees of freedom for the total sum of squares into fifteen sets of "additive" degrees of freedom, as for example

$\mathcal{L}$: (l - 1) d.f.

$\mathcal{F}$: (f - 1) d.f.

$\mathcal{L} \times \mathcal{F}$: (l -1)(f-1) d.f.

etc.

We denote mean squares according to the pattern used elsewhere in the thesis, for example, L* denotes the $\mathcal{L}$ mean square, I^*_{GH} denotes the $\mathcal{G}$ x $\mathcal{H}$ mean square and so on.

The expectations of these mean squares are easily written down using the patterns of Tables 35 and 36. We give the results, for the split split plot experiment outlined above, in Table 37. Table 38 gives implicit definitions of the Σ's used in Table 37 in terms of the components of variation of the population of conceptual true responses. To shorten Table 37 we use the following additional notation:

$$\Sigma_o = \sigma^2 + \Sigma_p + \Sigma_{pf} + \Sigma_{pg} + \Sigma_{ph} + \Sigma_{pfg} + \Sigma_{pfh} + \Sigma_{pgh} + \Sigma_{pfgh}$$

$$\Sigma_{oo} = h\Sigma_b + h\Sigma_{bf} + h\Sigma_{bg} + \Sigma_{bh} + h\Sigma_{bfg} + \Sigma_{bfh} + \Sigma_{bgh} + \Sigma_{bfgh}$$

$$\Sigma_{ooo} = gh\Sigma_s + gh\Sigma_{sf} + h\Sigma_{sg} + \Sigma_{sh} + h\Sigma_{sfg} + \Sigma_{sfh} + \Sigma_{sgh} + \Sigma_{sfgh}$$

$$\Sigma_{oooo} = fgh\Sigma_l + gh\Sigma_{lf} + h\Sigma_{lg} + \Sigma_{lh} + h\Sigma_{lfg} + \Sigma_{lfh} + \Sigma_{lgh} + \Sigma_{lfgh}$$

$$\Sigma^f_{oooo} = gh\Sigma_{lf} + h\Sigma_{lfg} + \Sigma_{lfh} + \Sigma_{lfgh}$$

$$\Sigma^g_{oooo} = h\Sigma_{lg} + h\Sigma_{lfg} + \Sigma_{lgh} + \Sigma_{lfgh}$$

$$\Sigma^h_{oooo} = \Sigma_{lh} + \Sigma_{lfh} + \Sigma_{lgh} + \Sigma_{lfgh}$$

$$\Sigma^g_{ooo} = h\Sigma_{sg} + h\Sigma_{sfg} + \Sigma_{sgh} + \Sigma_{sfgh}$$

Table 37. Expected mean squares for split split plot under general conditions.

Mean Square	Expected Mean Square
L^*	$\Sigma_{oooo} + \Sigma_{ooo} + \Sigma_{oo} + \Sigma_{o}$
F^*	$lgh\Sigma_{f} + lh\Sigma_{fg} + lg\Sigma_{fh} + l\Sigma_{fgh} + \Sigma^{f}_{ooo} + \Sigma_{ooo} + \Sigma_{oo} + \Sigma_{o}$
G^*	$lfh\Sigma_{g} + lh\Sigma_{fg} + lf\Sigma_{gh} + l\Sigma_{fgh} + \Sigma^{g}_{oooo} + \Sigma^{g}_{ooo} + \Sigma_{oo} + \Sigma_{o}$
H^*	$lfg\Sigma_{h} + lg\Sigma_{fh} + lf\Sigma_{gh} + l\Sigma_{fgh} + \Sigma^{h}_{oooo} + \Sigma^{h}_{ooo} + \Sigma^{h}_{oo} + \Sigma_{o}$
I^*_{LF}	$\Sigma^{f}_{oooo} + \Sigma_{ooo} + \Sigma_{oo} + \Sigma_{o}$
I^*_{LG}	$\Sigma^{g}_{oooo} + \Sigma^{g}_{ooo} + \Sigma_{oo} + \Sigma_{o}$
I^*_{LH}	$\Sigma^{h}_{oooo} + \Sigma^{h}_{ooo} + \Sigma^{h}_{oo} + \Sigma_{o}$
I^*_{FG}	$lh\Sigma_{fg} + l\Sigma_{fgh} + \Sigma^{fg}_{oooo} + \Sigma^{g}_{ooo} + \Sigma_{oo} + \Sigma_{o}$
I^*_{FH}	$lg\Sigma_{fh} + l\Sigma_{fgh} + \Sigma^{fh}_{oooo} + \Sigma^{h}_{ooo} + \Sigma^{h}_{oo} + \Sigma_{o}$

Table 37 (Continued)

I^*_{GH}	$1f\Sigma_{gh} + 1\Sigma_{fgh} + \Sigma^{gh}_{oooo} + \Sigma^{gh}_{ooo} + \Sigma^{h}_{oo} + \Sigma_{o}$
I^*_{LFG}	$\Sigma^{fg}_{oooo} + \Sigma^{g}_{ooo} + \Sigma_{oo} + \Sigma_{o}$
I^*_{LFH}	$\Sigma^{fh}_{oooo} + \Sigma^{h}_{ooo} + \Sigma^{h}_{oo} + \Sigma_{o}$
I^*_{LGH}	$\Sigma^{gh}_{oooo} + \Sigma^{gh}_{ooo} + \Sigma^{h}_{oo} + \Sigma_{o}$
I^*_{FGH}	$1\Sigma_{fgh} + \Sigma^{fgh}_{oooo} + \Sigma^{gh}_{ooo} + \Sigma^{h}_{oo} + \Sigma_{o}$
I^*_{LFGH}	$\Sigma^{fgh}_{oooo} + \Sigma^{gh}_{ooo} + \Sigma^{h}_{oo} + \Sigma_{o}$

Table 38. Implicit definitions of the Σ's in terms of population components of variation

---

$$\sigma_1^2 = \Sigma_1 + \frac{1}{F}\Sigma_{1f} + \frac{1}{G}\Sigma_{1g} + \frac{1}{H}\Sigma_{1h} + \frac{1}{FH}\Sigma_{1fh} + \frac{1}{GH}\Sigma_{1fg}$$
$$+ \frac{1}{FGH}\Sigma_{1fgh} + \frac{1}{S}\Sigma_s + \frac{1}{SF}\Sigma_{sf} + \frac{1}{SG}\Sigma_{sg} + \frac{1}{SH}\Sigma_{sh}$$
$$+ \frac{1}{SFH}\Sigma_{sfh} + \frac{1}{SFG}\Sigma_{sfg} + \frac{1}{SFGH}\Sigma_{sfgh} + \frac{1}{SB}\Sigma_b$$
$$+ \ldots\ldots + \frac{1}{SBFGH}\Sigma_{bfgh} + \frac{1}{SBP}\Sigma_p + \ldots\ldots$$
$$+ \frac{1}{SBPFGH}\Sigma_{pfgh}$$

$$\sigma_s^2 = \Sigma_s + \frac{1}{F}\Sigma_{sf} + \ldots\ldots + \frac{1}{FGH}\Sigma_{sfgh} + \frac{1}{B}\Sigma_b + \frac{1}{BF}\Sigma_{bf}$$
$$+ \ldots\ldots + \frac{1}{BFGH}\Sigma_{bfgh} + \frac{1}{BP}\Sigma_p + \ldots\ldots$$
$$+ \frac{1}{BPFGH}\Sigma_{pfgh}$$

$$\sigma_b^2 = \Sigma_b + \frac{1}{F}\Sigma_{bf} + \ldots\ldots + \frac{1}{FGH}\Sigma_{bfgh} + \frac{1}{P}\Sigma_p + \frac{1}{PF}\Sigma_{pf}$$
$$+ \ldots\ldots + \frac{1}{PFGH}\Sigma_{pfgh}$$

$$\sigma_p^2 = \Sigma_p + \frac{1}{F}\Sigma_{pf} + \ldots\ldots + \frac{1}{PFGH}\Sigma_{pfgh}$$

$$\sigma_f^2 = \Sigma_f + \frac{1}{G}\Sigma_{fg} + \frac{1}{H}\Sigma_{fh} + \frac{1}{GH}\Sigma_{fgh} + \frac{1}{L}\Sigma_{1f}$$
$$+ \frac{1}{LG}\Sigma_{1fg} + \frac{1}{LH}\Sigma_{1fh} + \frac{1}{LGH}\Sigma_{1fgh} + \frac{1}{LS}\Sigma_{sf}$$

Table 38 (Continued)

$$+ \frac{1}{LSG}\Sigma_{sfg} + \ldots. + \frac{1}{LSGH}\Sigma_{sfgh} + \frac{1}{LSB}\Sigma_{bf}$$

$$+ \ldots. + \frac{1}{LSBGH}\Sigma_{bfgh} + \frac{1}{LSBF}\Sigma_{pf} + \ldots.$$

$$+ \frac{1}{LSBPGH}\Sigma_{pfgh}$$

$$\sigma_g^2 = \Sigma_g + \frac{1}{F}\Sigma_{fg} + \frac{1}{H}\Sigma_{gh} + \ldots. + \frac{1}{LSBPFH}\Sigma_{pfgh}$$

$$\sigma_h^2 = \Sigma_h + \frac{1}{F}\Sigma_{fh} + \frac{1}{G}\Sigma_{gh} + \ldots. + \frac{1}{LSBPFG}\Sigma_{pfgh}$$

σ_{lf}^2 = Common Σ's of σ_l^2 and σ_f^2, using as coefficient the common part of the coefficients of the common Σ's.

σ_{lg}^2 = Common Σ's of σ_l^2 and σ_g^2, etc., as above.

σ_{lh}^2 = Common Σ's of σ_l^2 and σ_h^2, etc., as above.

The same rule applies to all other σ^2's. For example, to write σ_{bfgh}^2 we find the common parts of the right hand sides of σ_b^2, σ_f^2, σ_g^2, and σ_h^2. Thus

$$\sigma_{bfgh}^2 = \Sigma_{bfgh} + \frac{1}{P}\Sigma_{pfgh} .$$

$$\Sigma^{h}_{ooo} = \Sigma_{sh} + \Sigma_{sfh} + \Sigma_{sgh} + \Sigma_{sfgh}$$

$$\Sigma^{h}_{oo} = \Sigma_{bh} + \Sigma_{bfh} + \Sigma_{bgh} + \Sigma_{bfgh}$$

$$\Sigma^{fg}_{oooo} = h\Sigma_{lfg} + \Sigma_{lfgh}$$

$$\Sigma^{fh}_{oooo} = g\Sigma_{lfh} + \Sigma_{lfgh}$$

$$\Sigma^{gh}_{oooo} = f\Sigma_{lgh} + \Sigma_{lfgh}$$

$$\Sigma^{fgh}_{oooo} = \Sigma_{lfgh}$$

$$\Sigma^{gh}_{ooo} = \Sigma_{sgh} + \Sigma_{sfgh}$$

To show how replication within one of the experimental unit classifications would be handled we consider the case where, say, levels of g are replicated within sources, say each level appearing d times (i. e. on d blocks) within a source. The various results given in Table 37 would then need to be adjusted for the factor d according to the rule that each Σ component (i. e. reduced Σ component) has a coefficient equal to the number of observations for that classification. For example, Σ_l would have coefficient fghd, Σ_s would have coefficient ghd, Σ_b would have coefficient h, Σ_p would have coefficient 1, Σ_{sg} would have coefficient hd, etc.

Further, we would then have additional lines in the analysis of variance table: blocks within sources within location ($\mathcal{B}$), with lfg(d-1) degrees of freedom; $\mathcal{B} \times \mathcal{H}$ with lfg(d-1)(h-1) degrees of freedom. Then the expected mean square for $\mathcal{B}$ would be

$$E(B^*) = \Sigma_{oo} + \Sigma_o .$$

The expected mean square for $\mathcal{B} \times \mathcal{H}$ would be

$$E(I^*_{BH}) = \Sigma^h_{oo} + \Sigma_o .$$

If we had replicated $\mathcal{F}$ on sources within locations we would have: sources within locations ($\mathcal{S}$); $\mathcal{S} \times \mathcal{G}$; $\mathcal{S} \times \mathcal{H}$; $\mathcal{S} \times \mathcal{G} \times \mathcal{H}$, with expected mean squares

$$E(S^*) = \Sigma_{ooo} + \Sigma_{oo} + \Sigma_o \quad ;$$

$$E(I^*_{SG}) = \Sigma^g_{ooo} + \Sigma_{oo} + \Sigma_o \quad ;$$

$$E(I^*_{SH}) = \Sigma^h_{ooo} + \Sigma^h_{oo} + \Sigma_o \quad ;$$

$$E(I^*_{SGH}) = \Sigma^{gh}_{ooo} + \Sigma^h_{oo} + \Sigma_o \quad .$$

E. General Discussion and Extensions

The remarks of this division are to be taken as an integral part of the general developments reported on in Part II. Everything covered in this division is applicable to all the preceding divisions.

In referring to parts of the preceding divisions we shall reference as exemplified by: (A6) refers to Section 6 of Division A of Part II.

1. Basic assumptions

We summarize here the basic assumptions underlying the results of the preceding divisions. (More specialized assumptions are also used in special cases.) These are as follows:

(1) That there exist procedures by which we can make random selections and allocations.

(2) That the response from a given treatment on a given experimental unit does not depend on how other treatments are distributed over other units.

(3) That technical errors are in no way dependent on treatments (but see Case 6 of (B11) for an extended viewpoint).

(4) That technical errors act additively with respect to "true" responses for the given scale.

(5) That technical errors can be treated as uncorrelated random variables with means zero and constant variance.

Of these assumptions (1) is semi-philosophical; (2) seems to be necessary (at least approximately) if the notion of a treatment is to have much meaning; (3), (4) and (5) may be far from reality in some cases but will be reasonable often and will be unimportant if technical errors are small compared with unit errors.

Various modifications and generalizations of (3), (4) and (5) above are possible. One such modification of (3) is given under (B11), in which we sample individual "treatments" from within a "category". The category, or level, of $\mathcal{C}$ may be regarded as defining the treatment, and the sampling from within a category may be regarded as giving "treatment errors". No homogeneity assumptions are made with respect to "treatment errors within categories" in (B11).

Similarly the development of (B11) represents a generalization of assumption (4) above.

It is elementary to extend assumption (5) to the case where technical errors can be treated as uncorrelated random variables with the same (not necessarily zero) mean and with variances dependent (only) on the experimental units with which they are associated. We have not done so throughout simply to simplify writing.

A further level of generalization would be to the case where technical errors depended on the response. We have not given explicit consideration to such a generalization.

2. Randomization

The developments we have given are to a large extent hinged directly on procedures of random sampling and random allocation. The majority of the results given may therefore be regarded as having their validity and meaning because of randomization. Insofar as the additive part of unit errors is concerned these are completely controlled statistically by randomization in the sense that they do not introduce bias (on the average) into estimates and that the effect of their variability can be estimated. The interactive part of unit errors is not entirely controlled by randomization and does introduce systematic biases in the estimation of components of variation as we have seen; however linear estimates remain unbiased.

Fisher (1926, 1935a) has emphasized the importance of the physical act of randomization.* The point of view that a particular state of nature can be treated as though it were a random sample from some larger grouping may sometimes be convenient, but it would appear to be extremely hazardous, especially because of the absence of objectivity in such an evaluation. If such a construct of individual fancy were to be employed it would be very desirable to have its logical details made explicit. (See an example of such in Section 6 below.)

---

*We quote from Fisher (1926; pg. 508): "An experiment either admits a valid estimate of error, or it does not; whether it does so, or not, depends not on that actual arrangement of the plots, but only on the way that arrangement was arrived at".

It should however be made explicit that while randomization provides us with a safeguard against systematic bias in comparisons of operations we have actually carried out, it does not provide us with any procedure for identifying the essential characteristics or properties of the operations. This remains in the province of the experimenter's insight and intuition.

3. Advantages of the method

The method employed herein in the study of randomized experiments has a number of useful properties. The entire procedure is hinged around a single statistical model incorporating random variables, some of whose joint distributional properties are easily derived, and unknown population parameters, each of which have a physical interpretation. The model then summarizes all the relevant information about the physical situation, experimental design and procedure.

This model can be used quite formally to study statistical properties of interest such as expectations of estimates, variances of estimates, expectations of mean squares, estimation of error, etc., by purely algebraic procedures. There is no reason, in principle or practice, why the models could not be used, for example, to find variances and covariances of mean squares. In fact such models have been employed by Kempthorne (1952b) and by Wilk (1953a) to find the covariance structure of mean squares under null assumptions. It is true that some of the algebraic

procedures are tedious, but experience lightens the load; and there is always the hope that sufficient insight will lead to patterns of knowledge (as discussed below for the first moment of mean squares).

Perhaps the most important attributes of the method are that it lays bare the dependence of the meaning of "treatment effect" on the scale of observation; it leaves no ambiguity of the meaning of interaction and its dependence on the scale of observation; it brings out into the open the necessary assumptions (or conditions) for a meaningful interpretation of the analysis of variance; finally, and perhaps most important, it leads directly to the idea of the underlying functional structure of factors and units, imposing a (possible) recognizable) relation between certain components of the model. This latter point is discussed further in Section 7 below and in Part III of the thesis.

4. Relation of population and statistical model

We have not described heretofore the relation of the experimental design to the structure of the population model which was given. In fact, to simplify the presentation some aspects of the relation were hidden, in that certain interaction terms were not explicitly defined (though components of variation based on them are defined.) In a given situation, one's intuition is an adequate guide, but the procedure can be formalized to a considerable extent. We demonstrate the formalization for a two-factor completely randomized design, using the notation of (A 1, 2, 3).

The fundamental roles are played by the sampling and design random variables, the α_i^{i*}, β_j^{j*}, and ρ_m^{i*j*f}, whose properties effectively summarize what is done, and the conceptual true yields, the Y_{ijm}. It is easy to see that the observations, the x_{i*j*f}, are related to the Y_{ijm} by

$$x_{i*j*f} = \sum_{ijm} \alpha_i^{i*} \beta_j^{j*} \rho_m^{i*j*f} Y_{ijm},$$

where to simplify this discussion we ignore the technical errors, the ϵ_{ijm}.

Using the above relationship we now consider the breakdown of Y_{ijm} into means and deviations so that the random variables α_i^{i*}, β_j^{j*}, and ρ_m^{i*j*f} will appear none at a time, one at a time, all possible pairs, and all together.

Clearly only $Y_{...}$ will be independent of all three, since $\sum_i \alpha_i^{i*} = 1$, $\sum_j \beta_j^{j*} = 1$, $\sum_m \rho_m^{i*j*f} = 1$; only $Y_{i..}$ will be independent of β_j^{j*} and ρ_m^{i*j*f}, and so on. To balance the identity and maintain symmetry it is natural to define each population component so as to sum to zero on all of its subscripts. Thus we have

$$\begin{aligned} x_{i*j*f} = {} & Y_{...} + \sum_i \alpha_i^{i*} (Y_{i..} - Y_{...}) + \sum_j \beta_j^{j*} (Y_{.j.} - Y_{...}) \\ & + \sum_m \rho_m^{i*j*f} (Y_{..m} - Y_{...}) + \sum_{ij} \alpha_i^{i*} \beta_j^{j*} (Y_{ij.} - Y_{i..} - Y_{.j.} + Y_{...}) \\ & + \sum_{im} \alpha_i^{i*} \rho_m^{i*j*f} (Y_{i.m} - Y_{i..} - Y_{..m} + Y_{...}) \\ & + \sum_{jm} \beta_j^{j*} \rho_m^{i*j*f} (Y_{.jm} - Y_{.j.} - Y_{..m} + Y_{...}) \\ & + \sum_{ijm} \alpha_i^{i*} \beta_j^{j*} \rho_m^{i*j*f} (Y_{ijm} - Y_{ij.} - Y_{i.m} - Y_{.jm} + Y_{i..} + Y_{.j.} \\ & \qquad + Y_{..m} - Y_{...}) \end{aligned}$$

which gives us our statistical model and also indicates the appropriate definitions for the components of the population model. In the model given in (A3) we have amalgamated the last three components above to simplify writing.

The situation for the randomized block design is a little more complex than the above in that we have "nested" (or conditionally defined) random variables. Thus in Division B we specified the selection of treatments by β_k^{k*}, the selection of blocks by α_i^{i*}, and the random selection of units within selected blocks plus the randomization of selected treatments to these by ρ_{i*j}^{i*k*f}. Clearly here the ρ random variable has no interest by itself since it does not specify a population entity uniquely, and hence it must always be associated with α_i^{i*} which will identify the population block. The decomposition of the random variables which is used was chosen for the great convenience, in algebraic and statistical manipulation, of having the random variables be statistically independent.

Thus for the randomized block design we have

$$x_{i*k*f} = \sum_{ijk} \alpha_i^{i*} \beta_k^{k*} \rho_{i*j}^{i*k*f} Y_{ijk},$$

(again ignoring the technical errors ϵ_{ijk} for these discussions), and we now decompose Y_{ijk} into means and deviations according to the requirement that we have terms containing, respectively, none of α_i^{i*}, β_j^{j*}, ρ_{i*j}^{i*k*f}; α_i^{i*} only; β_k^{k*} only; $\alpha_i^{i*}\ \beta_k^{k*}$ only; $\alpha_i^{i*}\ \rho_{i*j}^{i*k*f}$ only; and all three. This gives

$$x_{i^*k^*f} = Y_{\ldots} + \sum_i \alpha_i^{i^*} (Y_{i..} - Y_{\ldots}) + \sum_k \beta_k^{k^*} (Y_{..k} - Y_{\ldots})$$

$$+ \sum_{ik} \alpha_i^{i^*} \beta_k^{k^*} (Y_{i.k} - Y_{i..} - Y_{..k} + Y\ldots)$$

$$+ \sum_{ij} \alpha_i^{i^*} \rho_{i^*j}^{i^*k^*f} (Y_{ij.} - Y_{i..})$$

$$+ \sum_{ijk} \alpha_i^{i^*} \beta_k^{k^*} \rho_{i^*j}^{i^*k^*f} (Y_{ijk} - Y_{ij.} - Y_{i.k} + Y_{i..})$$

For the latin square design we had two models, one for x_{k^*f} and one for $z_{i^*j^*}$. The model for $z_{i^*j^*}$ was based on the random variables $\alpha_i^{i^*}$, $\beta_j^{j^*}$, $\gamma_k^{k^*}$, $\rho_{i^*j^*}^{k^*}$, but we should, and did, combine $\gamma_k^{k^*}$ and $\rho_{i^*j^*}^{k^*}$ for the $z_{i^*j^*}$ model as $\sum_{k^*} \gamma_k^{k^*} \rho_{i^*j^*}^{k^*}$. Then we take all possible, singles, doubles and triples of these.

We see then that the prior definition of the population model is a matter of convenience in presentation and interpretation. The value of the entire breakdown into a linear model lies in the facts that the components of the model (and the corresponding components of variation) have a physical interpretation and are of interest, and also that the detailed breakdown allows one to appreciate more clearly the nature, meaning, and possible effect of any implicit or explicit assumptions in his procedure.

5. Functions estimable by mean squares

We have seen that, under general non-additivity conditions, there are characteristic linear combinations of defined components of variation (referred to as Σ's) which are estimable by the analysis of variance mean squares, whatever the relation of

population and sample sizes. The structure of the Σ's is of course dependent upon the particular definitions of the components of variation, which in turn depend upon the components of the population model. The preceding section shows however that the population model is essentially determined by the experimental procedure and design, which also determines the analysis of variance employed. Thus the chain of connection is complete, though it has a circular structure rather than a start and end since physical relationships influence the choice of the design and procedure.

We have not been able to attribute any physical significance to the Σ's but their unifying and simplifying effect is of considerable value. In particular the pattern of extension of results on expected mean squares becomes quite simple in terms of the Σ's.

In later sections we discuss the general implicit and explicit definition of the Σ's in terms of the defined components of variation. The Σ's are employed in giving detail (in Section 8 below) on a pattern of relevance in extensions of the results of the preceding divisions.

6. Effects of non-additivities

The importance of the effects of non-additivities is very much a matter of particular evaluation for a particular experiment.

We have seen, by some approximation arguments given in connection with Division B, that under some conditions interactions

of treatments with experimental units may be negligible if the units are structured into fairly homogeneous blocks. Similar arguments in Division C indicated that the three-way interaction of rows, columns and treatments may often be unimportant.

It is however worth making explicit the importance of the scale of measurement, insofar as non-additivities are concerned. If any classification is "physically homogeneous" then all interactions with that classification will be zero; but if the classification is not homogeneous then there will always exist scales of observation for which any designated interaction may be important. It will be apparent that the guidance of specific physical knowledge in a particular experimental situation is essential.

There is, in general, a "bias" in the analysis of variance due to "interactions with experimental units". The cause for this bias is the "fractional sampling" (referred to as random confounding below) arising from the fact that experimental units can be used only once. These interactions could be eliminated by choosing an appropriate (but unknown and often unknowable) scale for the observations. More realistically one can attempt to depress their importance for most situations by making experimental units as alike as economically feasible.

In fact, in most situations the bias of the analysis of variance (exemplified for example in the estimation of components of variation) due to unit-treatment interactions is made negligible by the factor $\frac{1}{\text{size of population of units}}$ which occurs. In some

situations it may be reasonable directly to take this factor as zero.

In other cases the following considerations may be of relevance. Suppose a designated bar of metal is cut up into P equal size pieces. One might consider the bar as consisting of R($>>$P) particles which have been randomly arranged into P groups of $Q = \frac{R}{P}$ particles each. There are $S = \prod_{i=0}^{P-1} \binom{R - iQ}{Q}$ possible distinct arrangements of the bar into P pieces. The units we use would then be regarded as a sample from the P available, while the entire arrangement would be regarded as a sample of one from the S possible. Further study of such a viewpoint might be instructive.

Many procedures of statistical experimental design are critically dependent upon implicit or explicit assumptions of additivity. (For example the interpretation of the latin square design.) These have undoubtedly proved very useful. But the cavalier disregard of interactions of factors or of treatments with groupings of experimental material will obscure important relations and conclusions under the guise of statistical objectivity.

The effect of non-additivities is closely connected with the subject of transformations. Some summary discussions, with references, of the purposes and procedures of transformations have been given by Rao (1952) and Kempthorne (1952a). The importance of transforming to attain additivity has been emphasized by Tukey (1949a, 1949b, 1955), Kempthorne (1952a), Kempthorne and

Barclay (1953), Moore and Tukey (1954) and Fisher (1954). Tukey (1949a) has given a test which is sensitive to non-additivity.

7. Error terms

We have seen the connection between error terms for tests, for estimation of components of variation, and for estimation of the average variance of estimates of differences of effects. But many aspects of the situation regarding error terms require considerable clarification.

There are difficult distributional theory problems in connection with both testing and estimation. Normality assumptions give "exact" results for simple situations. For more complex situations exact results are not available even under normal theory, though useful approximate procedures, based on normality assumptions, have been given or discussed by Satterthwaite (1946), Robbins and Pitman (1949), Cochran and Cox (1950), Cochran (1951) and Kempthorne (1952a). Again for simple situations, the effect of some departures from normality has been investigated (see some summary discussion by Davies (1954)). From the point of view of relieving the weight of statistical assumptions in significance tests the randomization test procedure (see for example Fisher (1936), Kempthorne (1952a)) is a reliable principle but the practical use of this principle would lie in finding appropriate approximations based on fitting moments, empirical investigations, and the suggestiveness of normal theory. Contributors in this

area have been indicated in the Introduction. Of relevance here too are the problems associated with "preliminary tests", which were first studied by Bancroft (1944). The relevant literature on this subject has been reviewed (and extended) in a recent report by Bozivich et al (1955) and Huntsberger (1955).

A reasonable type of approach on the distributional problems would be to superimpose, onto the derived model, a model which corresponded on certain important properties, but with sufficient (reasonable) homogeneity and other assumptions to enable "exact" treatment. Thus, for example, approximate tests for the mixed model case might be obtained using the covariance matrix of the components in the derived statistical model to define a multivariate normal distribution. Some references of relevance on this matter are Graybill (1954), Scheffé (1955). Such procedures are however subject to limitations (for example see page 150 Kempthorne (1952a)).

We have given considerable attention to the questions concerning the connection of the model to the experimental situation and design and the expectations of the analysis of variance mean squares, and have referred to other points of view, investigations, and results. This matter is in some ways fundamental, for the entire interpretation of the analysis of variance and of the experiment may be influenced by understandings of relationships among mean squares as indicated by their expectations. The tie-up with estimation is immediate in that everything flows from the understanding of the

statistical linear model. General results on variances and covariances of mean squares, based on derived linear models such as we have used above, would be valuable. Of relevance to this matter are Hooke's (1954a) results on variances and covariances of mean squares for "sampling from a matrix", valid for two-factor completely randomized experimental design situations under certain additivity assumptions.

There appears to be the point of view somewhat current in experimental statistics that it is the objective of an experiment which determines the appropriate statistical analysis rather than (or equally with) the sampling procedure and experimental design. Presumably a consequence of this might be that a physicist concerned with properties true for the entire solar system would be prepared to extend, by statistical (empirical) argument supposed to be based on probability considerations, experimental evidence obtained on Earth to all other planets simply by claiming that he will treat Earth "as though" it were a random selection from all planets of the system. We take the point of view that there is an important distinction between statistical inference and scientific inference (discussed somewhat by Wilk (1953a)) and that while the former is based on random sampling and probability, the latter depends on causal notions, logical approximations by mechanical pictures, etc. While we would regard it as entirely reasonable to treat an actual sample (randomly obtained) as fixed we cannot accept what would amount to the converse procedure. The

criticism of "quota sampling" in survey work carries over directly to experimental situations. The extension of a statistical inference to broader circumstances is certainly of all-important concern, but it is one which, we feel, does not fall directly in the province of statistics or of statisticians as such.

Ignoring the distributional problems, and assuming agreement on expected mean squares, we turn now to the question of what test ratios are meaningful to use in what circumstances. We have given some discussion of the influence of "choice of null hypothesis" on the selection of a test ratio in (B8). We shall not discuss further the possible interest in the test of "identity of treatments" (Fisher's (1935b) hypothesis) versus the test of "equality of average responses of treatments" (Neyman et al's (1935) hypothesis). However some remarks are in order as to the distinction between the selection and meaning of a test ratio for the fixed case as against that for the random case (to consider only the extremes).

Suppose we have a two-factor experiment in a completely randomized design, with both factors fixed under conditions of (A7). The conventional practice here would be to consider only the test ratios $\frac{I^*_{AB}}{R^*}$, $\frac{A^*}{R^*}$, $\frac{B^*}{R^*}$ for significance of $\mathcal{A}$ main effects, $\mathcal{B}$ main effects, and $\mathcal{A} \times \mathcal{B}$ interactions. It seems, however, that even for the fixed case one might have interest in showing the statistical significance for the possible excess of, say, $\mathcal{A}$ main effects over $\mathcal{A} \times \mathcal{B}$ interactions. Of course distributional theory is a big stumbling block, insofar as the analysis of variance is concerned,

and it may be necessary to use auxilliary approaches for further objective evaluation. A relevant reference is Cochran (1951). But from a superficial point of view one can simply "look at the mean squares".

Let us consider now the case when $\mathcal{A}$ and $\mathcal{B}$ are both random, again under conditions of (A7). The expectations of the mean squares are now such that both main effect mean squares contain, on the average, contributions due to interactions. This is a consequence of our wider area of interest (justified statistically by the procedure of random sampling) and of considerable relevance so far as estimation of components of variation attributable to main effects and to interactions is concerned. But insofar as the actual mean squares are concerned, each of A*, B* and I^*_{AB} are determined by the same data. Thus even though estimates of σ^2_a, σ^2_b and σ^2_{ab}, defined for $\mathcal{A}$ and $\mathcal{B}$ random, are meaningfully obtained from the analysis of variance mean squares, these estimates will in fact be functionally related as well as statistically dependent. What we would really like to have in a statistical test of significance of the null hypothesis, say $\sigma^2_a = 0$ in the random case, is a mean square whose expectation is the same as that of I^*_{AB}, but which is based on observations distinct from those used in A*. For as we shall see in Part III, the underlying functional structure relationships of factors to the observation will determine a functional relationship between the interaction components of the model and the main effect components of the model, and consequently

between the estimates of these and between any functions of the estimates of these. We know of no relevant work on this matter aside from the suggestiveness of Part III of this present thesis, but it would seem that the problem would be an important and fruitful one from the point of view of experimental statistics.

8. Patterns for expected mean squares

The object of this section is to discuss informally the extension of the given results on expected mean squares. The extensions are implicit in the pattern of these results but a discussion of the overall picture may be helpful. We confine attention to the case of orthogonal analyses with equal numbers in the "cells".

A general discussion of the analysis of variance is difficult, maybe impossible. The structure of the analysis is of course determined by the design of the experiment, which in turn is conditioned by the experimental situation. We have seen above the connection between the statistical model and the population model, that is, how the design and procedure implies the definition of the various population parameters. We shall take for granted then the "proper" analysis and the "proper" (corresponding) definitions of the various components of variation.

a. Types of components of variation. We abbreviate "components of variation" as "cv". There are various types of cv which we now list, develop notation for, and exemplify with reference to the division and section. We use α, β, γ, ... as "dummy" latin letters.

(i) Main effect unnested cv: (M_a).

Examples: (A10), $\sigma_a^2 = (M_a)$

(B5), $\sigma_b^2 = (M_b)$

(C5), $\sigma_r^2 = (M_r)$

(D4), $\sigma_h^2 = (M_h)$.

(ii) Main effect, first-level nested, cv: (M_β^a).

Examples: (A), none

(B5), $\sigma_p^2 = (M_p^b)$

(B11), $\sigma_t^2 = (M_t^c)$

(C), none

(D4), $\sigma_b^2 = (M_b^s)$

(iii) Main effect, second-level nested, cv: $(M_\gamma^{\alpha\beta})$.

(iv) Etc.

(v) Two category interactions: $(M_a) \times (M_{a'})$, or $(M_{a'}) \times (M_\beta^a)$, or $(M_{\beta'}^{a'}) \times (M_\beta^a)$, etc.

Examples: (A10), $\sigma_{bc}^2 = (M_b) \times (M_c)$

(B5), $\sigma_{tp}^2 = (M_t) \times (M_p^b)$

(B11), $\sigma_{pt}^2 = (M_p^b) \times (M_t^c)$

(C5), $\sigma_{rt}^2 = (M_r) \times (M_t)$

(vi) Three category interactions: $(M_a) \times (M_{a'}) \times (M_{a''})$, or $(M_{a''}) \times (M_{\beta'}^{a'}) \times (M_\gamma^{\alpha\beta})$, etc.

Examples: (A10), $\sigma_{abc}^2 = (M_a) \times (M_b) \times (M_c)$

(B9), $\sigma_{ghp}^2 = (M_g) \times (M_h) \times (M_p^b)$

(C5), $\sigma_n^2 = \sigma_{rct}^2 = (M_r) \times (M_c) \times (M_t)$.

(vii) Etc.

b. Definition of the Σ's. We shall use the notation of (a) above and classify the Σ's as corresponding to (M_α), or (M^α_β) or $(M_\alpha) \times (M_{\alpha'})$ or $(M_{\alpha''}) \times (M_{\alpha'}) \times (M^{\alpha\beta}_\gamma)$, etc.

Let (S_α) = size of population corresponding to an unnested classification; e.g. (A6): $(S_b) = B$, $(S_p) = P$; (B5): $(S_t) = T$.

Let (S^α_β) = size of the nested population corresponding to β (within α); e.g. (B5): $(S^b_p) = P$; (D4): $(S^s_b) = B$.

Similarly define $(S^{\alpha\beta}_\gamma)$; e.g. (D4): $(S^{sb}_p) = P$; and so on.

We now define "products" of the (M)'s according to the type of rule exemplified by

$$(M_\alpha) \times (M_\alpha) = (M_\alpha)$$
$$(M^\alpha_\beta) \times (M_\alpha) = (M^\alpha_\beta)$$
etc.

The types and definitions of the Σ's are as follows:

(i) Σ corresponding to (M_α), where α has nothing nested in it. Let $\alpha_1, \alpha_2, \ldots, \alpha_n$ denote all unnested main categories which are "crossed" with α. Then

$$\Sigma_\alpha = (M_\alpha)\left[1 - \frac{1}{(S_{\alpha_1})}(M_{\alpha_1})\right]\left[1 - \frac{1}{(S_{\alpha_2})}(M_{\alpha_2})\right] \cdots \left[1 - \frac{1}{(S_{\alpha_\beta})}(M_{\alpha_n})\right],$$

where the product is formally expanded and the products such as $(M_\alpha) \times (M_{\alpha_1})$ interpreted as the corresponding cv.

(ii) Σ corresponding to (M_α) where β is nested in α (and possibly nested in β, etc.) and where $\alpha_1, \alpha_2, \ldots, \alpha_n$ are main unnested categories crossed with α. Then

$$\Sigma_a = (M_a)\left[1 - \frac{1}{(S^a_\beta)}(M^a_\beta)\right]\left[1 - \frac{1}{(S_{a_1})}(M_{a_1})\right] \cdots \left[1 - \frac{1}{(S_{a_n})}(M_{a_n})\right],$$

where the product is expanded formally and interpreted, where necessary according to the given rules, as cv. Note that the subnested categories (of β) do not enter, the first-nested category of a acts like a cross, and any nested categories with $a_1, \ldots, a_n$ do not enter.

Example, (B5),

$$\Sigma_b = \sigma^2_b - \frac{1}{T}\sigma^2_{bt} - \frac{1}{P}\sigma^2_p + \frac{1}{TP}\Omega^2_{tp}$$

$$= (M_b) - \frac{1}{(S_t)}(M_b) \times (M_t) - \frac{1}{(S^b_p)}(M^b_p) + \frac{1}{(S_t)(S^b_p)}(M_t) \times (M^b_p)$$

$$= (M_b)\left[1 - \frac{1}{(S^b_p)}(M^b_p)\right]\left[1 - \frac{1}{(S_t)}(M_t)\right].$$

(iii) Σ corresponding to (M^a_β): It is handled just like a main category Σ, ignoring the a category.

(iv) Σ corresponding to $(M_a) \times (M_{a'})$, where a and a' are main categories with nothing nested in them. Let $a_1, \ldots, a_n$ be main categories crossed with both a and a'. Then

$$\Sigma_{aa'} = (M_a) \times (M_{a'}) \prod_{i=1}^{n}\left[1 - \frac{1}{(S_{a_i})}(M_{a_i})\right].$$

(v) Σ corresponding to $(M_a) \times (M_{a'})$ where a has β nested in it, a' has β' nested in it (and β and β' may have further hierarchal structure), and $a_1, \ldots, a_n$ are main categories crossed with a and a'.

$$\Sigma_{aa'} = (M_a) \times (M_{a'}) \left[1 - \frac{1}{(S^a_\beta)}(M^a_\beta)\right]\left[1 - \frac{1}{(S^{a'}_{\beta'})}(M^{a'}_{\beta'})\right]\left[1 - \frac{1}{(S_{a_1})}(M_{a_1})\right]$$

$$\cdots\cdots\cdots \left[1 - \frac{1}{(S_{a_n})}(M_{a_n})\right] .$$

Example, B(11), with n = 0,

$$\Sigma_{bc} = \sigma^2_{bc} - \frac{1}{T}\sigma^2_{bt} - \frac{1}{P}\sigma^2_{pc} + \frac{1}{TP}\sigma^2_{pt}$$

$$= (M_b) \times (M_c)\left[1 - \frac{1}{(S^c_t)}(M^c_t)\right]\left[1 - \frac{1}{(S^b_p)}(M^b_p)\right] .$$

(vi) Other definitions follow the identical pattern.

Example, (D4),

$$\Sigma_{ghb} = \sigma^2_{ghb} - \frac{1}{P}\Omega^2_{ghp}$$

$$= (M_g) \times (M_h) \times (M^s_b)\left[1 - \frac{1}{(S^{sb}_p)}(M^{sb}_p)\right].$$

c. Relation of σ^2 and Ω^2 to Σ's. First we note that there is really no distinction between the σ^2 and the Ω^2. The definitions of each are in general on the same basis, i.e. a sum of squares divided by (number of squares - number of linear dependencies). The point of the two notations was that Ω^2 would be used for those components which might reasonably be supposed small*. In what follows we write about σ^2's, but imply both σ^2's and Ω^2's.

*This has not however been entirely consistent throughout the thesis.

It is simple to express a σ^2 quantity in terms of Σ's. Suppose we want σ^2_a where β and γ nest in a, a' crosses with a and no other categories exist. Then

$$\sigma^2_a = \Sigma_a + \frac{1}{(S^a_\beta)}\Sigma_\beta + \frac{1}{(S^a_\beta)(S^{a\beta}_\gamma)}\Sigma_\gamma + \frac{1}{(S_{a'})}\Sigma_{aa'}$$

$$+ \frac{1}{(S^a_\beta)(S_{a'})}\Sigma_{\beta a'} + \frac{1}{(S^a_\beta)(S^{a\beta}_\gamma)(S_{a'})}\Sigma_{\gamma a'} \quad .$$

In general if a nests β nests γ etc. and a', a'', etc. cross with a, and a' nests β' nests γ' etc., and a'' nests β'' nests γ'', etc., and so on, then with the proper definition of "products" for the (Σ)'s we have

$$\sigma^2_a = (\Sigma_a)\left[1 + \frac{1}{(S^a_\beta)}(\Sigma^a_\beta) + \frac{1}{(S^a_\beta)(S^{a\beta}_\gamma)}(\Sigma^{a\beta}_\gamma) + \ldots\right]$$

$$\times \left[1 + \frac{1}{(S_{a'})}(\Sigma_{a'}) + \frac{1}{(S_{a'})(S^{a'}_{\beta'})}(\Sigma^{a'}_{\beta'}) + \ldots\right]$$

$$\times \left[1 + \frac{1}{(S_{a''})}(\Sigma_{a''}) + \frac{1}{(S_{a''})(S^{a''}_{\beta''})}(\Sigma^{a''}_{\beta''}) + \ldots\right]$$

$$\times \ldots\ldots\ldots,$$

where

$(\Sigma_a) \times (\Sigma^a_\beta) = \Sigma_\beta$, ($\beta$ nests within a),

$(\Sigma_a) \times (\Sigma_{a'}) = \Sigma_{aa'}$,

$(\Sigma_a) \times (\Sigma^a_\beta) \times (\Sigma^{a'}_{\beta'}) = \Sigma_{\beta\beta'}$, ($\beta$ nests in a, β' nests in a'),

etc.

Example, (C5),

$$\sigma_t^2 = (\Sigma_t)\left[1 + \frac{1}{R}(\Sigma_r)\right]\left[1 + \frac{1}{C}(\Sigma_c)\right].$$

Example, (D4),

$$\sigma_s^2 = (\Sigma_s)\left[1 + \frac{1}{B}(\Sigma_b^s) + \frac{1}{BP}(\Sigma_p^{sb})\right]\left[1 + \frac{1}{G}(\Sigma_g)\right]\left[1 + \frac{1}{H}(\Sigma_h)\right].$$

d. Coefficients of Σ's in expected mean squares. The coefficient of a Σ in the expected mean square for a balanced situation, is the number of observations 'contained in' the category designated in the experiment.

For example, (B5), the number of observations per block is rt, which is the coefficient of Σ_b.

For example, (C5), the number of observations per row-column intersection is 1, which is the coefficient of Σ_{rc}; the number of observations per row is t, which is the coefficient of Σ_r.

For example, (D4), the number of observations per $\mathcal{H}$ -block category is r, which is coefficient of Σ_{hb}.

In the next section we shall refer only to Σ quantities, the inclusion of the appropriate coefficient being understood.

e. Expected mean squares (ems). We shall indicate some general patterns concerning analysis of variance ems, in terms of Σ's. The Σ's depend only on population quantities; and their explicit and implicit definition in terms of the σ^{2}'s and Q^{2}'s has been discussed under (b) and (c).

The coefficients of the Σ's in the ems depend on sample sizes according to the pattern described under (d). To simplify writing, no explicit mention will be made of the (sample) coefficient that accompanies the Σ quantity in an ems.

Some background is required on classification relationships in the experiment. Since the experimental design and procedure in fact determines the population as well as the structure of the observations, the whole matter can be formalized into "types of sampling relationships" based on the concept of "true" yield. We shall however attempt to maintain the framework (and language) of the actual experiment and of how the data are classified.

The three types of relationships are crossed, hierarchal, and "randomly confounded" (or "randomized over", or "randomly fractionated"). We shall illustrate the use of these terms, sufficient for the content of the sequel.

Refer to (A1): $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{C}$ are jointly crossed, nothing is hierarchal, factors are randomly confounded (randomized over) with experimental units.

Refer to (B1): Treatments are crossed with blocks (since, for selected entities, every treatment appears in every block and every block contains every treatment), experimental units are hierarchal within blocks, treatments are randomly confounded with units-within-blocks (randomized over selected units within selected blocks).

Refer to (B11): $\mathcal{C}$ is crossed with blocks, treatments are hierarchal within $\mathcal{C}$, units are hierarchal within blocks, $\mathcal{C}$ is randomly

confounded with units-within-blocks, treatments-within-$\mathcal{C}$ are randomly confounded with blocks, treatments-within-$\mathcal{C}$ are randomly confounded with units-within-blocks.

Refer to (C1): Rows, columns and treatments are mutually crossed, rows are randomly confounded with column-treatment interactions, columns are randomly confounded with row-treatment interactions, and treatments are randomly confounded with row-column interactions. (The latin square design guarantees an observation for every row-treatment, (column-treatment, row-column) category, but the column (row, treatment) which is associated with a given row-treatment (column-treatment, row-column) is determined by a random process.)

Refer to (D1): Experimental units are hierarchal within blocks within sources, $\mathcal{G}$ and $\mathcal{H}$ and sources are jointly crossed (since there will be observations corresponding to every combination of (selected levels of) $\mathcal{G}$, $\mathcal{H}$ and sources), $\mathcal{G}$ is randomly confounded with blocks-within-sources, $\mathcal{H}$ is crossed with blocks, $\mathcal{G}$ is crossed with units-within-(randomly confounded) blocks, $\mathcal{H}$ is randomly confounded with units-within-blocks, (because for every existent $\mathcal{G}$-block category there is an observation corresponding to every $\mathcal{G}$-unit-within-block category).

We shall describe the pattern for ems for "main effects" in terms of the relationships described above; and then indicate the simple way of obtaining expected mean squares for "interactions". We recall that there are nested as well as unnested main effects mean squares; for example (D4, Table 31) we have the unnested main effects for $\mathcal{A}$, $\mathcal{G}$, $\mathcal{H}$, and the nested main effects for $\mathcal{B}$ (within $\mathcal{A}$) and for $\mathcal{P}$ (within $\mathcal{B}$ within $\mathcal{A}$).

Let a denote a main effect category. The ems corresponding to the a mean square is constructed as follows:

(i) It will contain main effect Σ components (1) for a, (2) for every category nested with a, (3) for every main category (perhaps hierarchal) randomly confounded with a, (4) for every hierarchal category of (3).

(ii) It will contain two-way interaction Σ components (1) for all two-way interactions involving any of the main effect components in (i), (2) for all additional two-way interactions randomly confounded with a, not involving a.

(iii) It will contain three-way interaction components along the same lines as (ii), and so on.

(iv) Delete all repetitions of Σ components.

(v) Put in appropriate coefficients (see (d) above).

We give an illustration of this from (B11): We have for the ems for blocks

(i) (1) Σ_b, (2) Σ_p, (3) Σ_t, (4) none

(ii) (1) Σ_{bc}, Σ_{bt}, Σ_{pc}, Σ_{pt}, Σ_{bt}, Σ_{pt}, (2) none

(iii) None.

Deleting repetitions and putting the appropriate coefficients we get the result given in Table 26.

As another illustration, refer to (D7). The listing for the ems for $\mathfrak{g}$ is

(i) (1) Σ_g, (2) none, (3) Σ_b, (4) Σ_p

(ii) (1) Σ_{gs}, Σ_{gh}, Σ_{gb}, Σ_{gp}, Σ_{gb}, Σ_{hb}, Σ_{gp}, Σ_{hp}, (2) none

(iii) (1) Σ_{ghs}, Σ_{ghb}, Σ_{ghp} .

Deleting repetitions, etc., we get the result of Table 35.

The procedure for a nested main effect category is the same; one simply ignores the "mother" category.

The procedure for most interactions is simple once the main effects ems are written down. One simply takes the "product" of the main effects ems, where by product we mean "the common part". For exceptions to this, use the same general procedure as for main effect ems.

We illustrate further using the latin square situation. The ems for rows contains

(i) (1) Σ_r, (2) none, (3) none;

(ii) (1) Σ_{rc}, Σ_{rt}, (2) Σ_{ct}.

(iii) (1) Σ_{rct}.

The "discrepance" can be regarded as the row-treatment interaction, which is randomly confounded with the column-treatment and row-column interactions, whence the ems for D* contains

(i) (1) Σ_{rt}, (2) none, (3) Σ_{ct}, Σ_{rc} .

(ii) (1) Σ_{rct} .

III. LINEAR MODELS AND FUNCTIONAL STRUCTURE

A. Introductory Remarks

The object of this part of the thesis is to report some preliminary progress on an investigation of functional structure by statistical techniques using defined linear statistical models. A review of the relevant literature is given in Division B. Preliminary original work is reported in Division C.

In Division E of the Introduction (Part I) we have given some introductory discussion on models, and in particular have referred to descriptive models and definitional models. It was pointed out there that particular descriptive models correspond to types of functional structure or classes of functions. Definitional models have been examined at some length in Part II, and the consequences of various types of non-additivities detailed.

Further remarks will be restricted to cases involving two factors $\mathcal{A}$ and $\mathcal{B}$. It will be assumed that the levels of these factors may be treated as though they were defined continuously, though the particular levels which we examine will be taken as a fixed finite set. We assume that there exist quantitative characteristics which define the levels of $\mathcal{A}$ and $\mathcal{B}$, denoted by u and v respectively. No direct knowledge concerning the nature or magnitude of the variables u and v is assumed, simply their

existence. Often the variables u and v will be complex functions of observable characteristics of the levels of the factors.

We use the term "functional structure" in the following sense: Two functions, each of two variables, $g(u, v)$, $h(r, s)$, have the same functional structure if there exist functions

$$\alpha = \alpha(u), \quad \beta = \beta(v)$$

such that

$$g(u, v) = h(\alpha, \beta).$$

If we look on the descriptive model as being the "simplest" or "most elementary" function of the class of functions all having the same functional structure (and assuming it is possible to agree on what is "simplest") then evidently two functions have the same functional structure if they have the same descriptive model. For example the functions

$$c_1x^2 + c_2 \log y, \quad \sin x + \cos y, \quad e^x + \int_0^y e^{-t^2}dt, \quad \alpha + \beta$$

all have the same functional structure, and it seems reasonable to take $(\alpha + \beta)$ as the descriptive model for this class of functions. The additive descriptive model has of course received a great deal of attention and use in statistics. By and large it is the only one which has been examined or applied generally. There has been recognition that the additive descriptive model may not be adequate everywhere. In a two factor completely randomized experiment with treatment combinations replicated the comparison of interaction mean square with residual mean square is a test of the

adequacy of the additive descriptive model. For a situation in which two such mean squares are not available, as for example the usual randomized block experiment with each treatment occurring once in each block, Tukey (1949a) has given a test sensitive to non-additivity, based on isolating one degree of freedom from the interaction sum of squares.

The question arises as to what one should do when non-additivity (i. e. inadequacy of the descriptive additive model) has been demonstrated. One answer is to look for a transformation of the observations to a scale such that additivity does hold. However the proper transformation to make is usually unknown and would in general involve unknown parameters. Thus many transformations would, in general, need to be examined. A bold approach of this type was used by Moore and Tukey (1954).

Another approach would be to find "estimated expected values" based on a postulated functional structure and examine deviations of observed values from these on the original scale. This is essentially the approach of Fisher and Mackenzie (1923) and Balmukand (1928).

A third approach is illustrated in Division C below. It is based on the fact that the functional structure imposes certain restrictions and relations among the parameters of the (population) model.

B. Review of Relevant Literature

We give in this division a brief outline of some literature of relevance to the problem of investigation of functional structure by statistical techniques.

The earliest reference we know of is Fisher and Mackenzie (1923). They employed the analysis of variance to evaluate a "product formula" relating yield to the product of two factors, one depending on the variety and the other on manurial treatment, and described an iterative procedure for finding the least squares solutions in fitting the formula. They had available an "error mean square" based on replication of treatment combinations. The sum of squares for treatment combinations was decomposed into three components attributable to deviations among computed values, difference between computed value mean and observed mean and deviations from the product formula, respectively. The mean square derived from the latter sum of squares they compared with the error mean square.

Balmukand (1928) used the analysis of variance to examine the adequacy of a "resistance formula" for yield-factor relations. His technique involved a transformation to reciprocals ($\frac{1}{y}$) which, under the resistance formula, were supposed to have an additive structure. He used as a first approximation to the reciprocal of the "expected" yield ($\frac{1}{m}$) estimates based on marginal means of the transformed yields. He then used as weights the fourth power of these crude

expectations to obtain an improved fit and then considered additional refinements. After inverting these he calculated the sum of squares, $\Sigma(y-m)^2$, and compared the mean square so obtained with a mean square derived from the interaction of blocks of experimental units with treatment combinations.

As part of a paper of considerable scope Tukey (1949b) gave an extensive discussion on "choice of terms". He pointed out the advantages which may be attained by studying data on an appropriately transformed scale - advantages both in basic understanding and in efficiency of estimation - and discussed the reduction of interaction in a "row-by-column design", restricting his treatment to linear combinations

$$y = af(x) + bg(x)$$

of two functions f and g, which may either be suggested by some experience or will need to be guessed. His suggested procedure is to choose the ratio of a to b so as to maximize the ratio

$$\frac{SSRy + SSCy}{SSBy}$$

where B= balance, R=row, C=column.

Tukey (1949a) described a statistical test procedure, based on isolating one degree of freedom in the analysis of variance, sensitive to non-additivity of classifications. Tukey justifies his choice of the one degree of freedom by considering a function

$$f\pi(a_{ij}) = a_{ij} + \pi(a_{ij} - a)^2$$

where π is a small constant,

$$a_{ij} = x_i + y_j \text{ (i.e. } a_{ij} \text{ exhibits perfect additivity),}$$

and a is taken as $\bar{x} + \bar{y}$.

He decomposes $f(a_{ij})$ into three components, two depending on column and row classifications alone, respectively, and $\pi (x_i - \bar{x})(y_i - \bar{y})$ which "contains all the non-additive effect due to analysis in terms of f(a) instead of in terms of a". He then argues that the difference between the i^{th} column mean and the grand mean is

$$(x_i - \bar{x}) + \pi\left((x_i - \bar{x})^2 - \overline{(x_i - \bar{x})^2}\right)$$

which is nearly $(x_i - \bar{x})$ when π is small and hence suggests isolating the degree of freedom indicated by the coefficients

(col. mean - grand mean)(row mean - grand mean).

His paper contains also some discussion of what to do (and not to do) if significant non-additivity is observed.

Ward and Dick (1952) considered the case when effects of blocks and treatments may be represented multiplicatively as

$$y_{ij} = (m' + b'_j + \epsilon'_{ij})(m'' + t''_i + \epsilon''_{ij}),$$

which they put in the form

$$y_{ij} = m + t_i + b_j + \mu t_i b_j + \epsilon_{ij}$$

and then gave an iterative solution for the estimates of the parameters from minimizing $\Sigma \epsilon^2_{ij}$. Their method appears to resemble that of Fisher and Mackenzie (1923) of whose work they were apparently not aware. They pointed out the relation of the reduction in sum of squares from their procedure to that of Tukey's (1949a).

Moore and Tukey (1954) apply Tukey's (1949a) test for non-additivity in answering a query (in Biometrics). Finding significant non-additivity they consider a two parameter family of transformations $(x + c)^p$ and propose a "95% confidence region for the additive transformation" (if it does in fact belong to this family). As auxiliary information they calculate and plot the ratio of mean square for treatments to mean square for interaction for the various transforms.

Anscombe and Tukey (1954) reviewed and proposed various methods of examining and testing data for non-additivity (and also for non-constancy of variance and non-normality), including some rough graphical procedures as well as more refined techniques.

Tukey (1955) illustrated how to apply his (1949a) test for non-additivity to a latin square and also discussed the general structure of the test.

C. A Test for a Multiplicative Functional Structure

In this division we present, informally, some preliminary ideas and results on examining a set of data, based on a two factor experiment, for multiplicative functional structure on a given scale.

1. The situation under consideration

The situation we shall consider is a two factor experiment under conditions as described in Division A of this Part III. Further, we suppose that a specific levels of $\mathcal{A}$ and b specific

levels of $\mathcal{B}$ are selected and the ab selected treatment combinations are studied in a completely randomized experiment involving r replicates of each combination, using rab experimental units selected at random from a population of size P. Generalization of the notions given here to some other situations is immediate.

It has been mentioned in (IB1) that it will often be of sufficient concern to know the relation of the factors with respect to average (over all experimental units) response. Furthermore by the very structure of the experiment, in randomizing over units, it is implicit that our interest is (or must be) directed to average behaviour of the factors.

The full definitional statistical model appropriate here is a specialization of the one given as Case 5 of the completely randomized design in (IIA9). We shall however restrict ourselves at the outset with the assumption of unit-treatment additivity (and will further strengthen our assumptions later to help obtain additional insight). The initial model we consider then is that of (IIA7), which for the present conditions becomes

$$x_{ijf} = \mu + a_i + b_j + (ab)_{ij} + \sum_m \rho_m^{ijf}(p_m + \epsilon_{ijm})$$

$$= \mu + a_i + b_j + (ab)_{ij} + e_{ijf}$$

where the components of the model have the definitions and interpretations given in (IIA2, 3) and we have put $e_{ijf} = \sum_m \rho_m^{ijf}(p_m + \epsilon_{ijm})$ to simplify writing.

We shall use the symbols u and v to denote quantitative characteristics (type and magnitude not necessarily known) which define,

so far as responses are concerned, the level of $\mathcal{A}$ and $\mathcal{B}$, respectively. These are identified for the levels used in the experiment by writing, for example, u_i to correspond to the i^{th} level of $\mathcal{A}$.

Let $Y_{ij.}$, the "true" response from treatment combination (i, j) for a given scale of observation averaged over all experimental units, be related to u_i and v_j by

$$Y_{ij.} = g(u_i, v_j)$$

where g is a function such that

$$\bar{Y} = g(u, v)$$

for any levels of $\mathcal{A}$ and $\mathcal{B}$ covered by range of the selected levels, and $\bar{Y}$ is the average response over all units.

2. The postulated descriptive model

We shall concern ourself with the problem of examining the data with respect to the adequacy of a multiplicate descriptive model. More specifically, suppose that $g(u, v)$ has a multiplicative-type functional structure so that there exist function

$$\alpha = \alpha(u), \quad \beta = \beta(v),$$

such that for all levels of $\mathcal{A}$ and $\mathcal{B}$, over the range defined by the respective a and b selections, the average true response is given by

$$\bar{Y} = g(u, v) = k + \alpha\beta$$

where k is an unknown constant.

It is of interest to note that this general multiplicative descriptive model depends on an unknown constant, while the general additive descriptive model, namely $(\alpha + \beta)$, does not involve any irremovable unknown constants.

For the particular levels examined in the experiment we would have

$$\alpha_i = \alpha(u_i),\ \beta_j = \beta(v_j)$$

and under the postulated functional structure

$$Y_{ij.} = g(u_i, v_j)$$

$$= k + \alpha_i\beta_j .$$

As something of an aside we show that

$$h(r, s) = k_1 + k_2 r + k_3 s + k_4 rs$$

is equivalent (in the obvious sense of the context) to $(k + \alpha\beta)$. To see this one needs only to put

$$\alpha = r + 1/k_4$$

$$\beta = k_3 k_4 s + k_2.$$

3. Possible approaches

To examine the adequacy of a multiplicative functional structure, or to test for specific values of k, or to estimate k, we might use an approach based on transforming to additivity.

Under the postulated descriptive model for $Y_{ij.}$ we would have, formally,

$$Z_{ij.} = \log (Y_{ij.} - k) = (\log \alpha_i) + (\log \beta_j) ,$$

and Z_{ij} would have an additive functional structure. Thus to test the plausibility of the multiplicative structure for $Y_{ij.}$ we might try to find a value for k, if such existed, such that in the analysis of

variance based on

$$z_{ijf} = \log (x_{ijf} - k)$$

the interaction mean square is not "too large" as compared with the residual mean square (which will exist if $r > 1$). If such a test was in the affirmative then we might map an interval estimate for k based on non-significance of corresponding variance ratios. Such a procedure would be along the lines of that used by Moore and Tukey (1954).

Another approach, similar in principle but rather different in procedure to the one above, would be to obtain "predicted" values of $z_{ij.}$ based on

$$m_{ij} = (z_{i..} + z_{.j.} - z_{...})$$

for a range of values of k. For any given value of k, we then transform back to find the "predicted" value of $x_{ij.}$ as

$$n_{ij} = e^{m_{ij}} + k .$$

Then the adequacy of the multiplicative descriptive model for that particular value of k is assessed by comparing the mean square based on

$$r\sum_{ij} (x_{ij.} - n_{ij})^2$$

with the residual mean square based on

$$\sum_{ijf} (x_{ijf} - x_{ij.})^2$$

The extension to finding an interval estimate for k by this method is immediate. This appears to be, essentially, the idea behind the

procedures of Fisher and Mackenzie (1922) and Balmukand (1928).

Aside from distributional questions (which are always with us) associated with either of the above methods there are the problems of heavy computation (probably heavier for the first procedure than the second) and the lack of any rough and simple preliminary evaluation procedure. Of course, with respect to computation, what is being done is rather ambitious and something must be paid out for this.

4. The basis of the present approach

The basis of the present approach is an examination of the relations and properties of components of the definitional statistical model which would be a consequence of a particular type of functional structure. Thus, for example, a consequence of an additive functional structure would be that

$$(ab)_{ij} = Y_{ij.} - Y_{i..} - Y_{.j.} + Y_{...}$$

would be zero, for every i and j.

If the postulated multiplicative descriptive model is correct then

$$Y_{ij.} = k + \alpha_i\beta_j$$

$$Y_{i..} = k + \alpha_i\beta_.$$

$$Y_{.j.} = k + \alpha_.\beta_j$$

$$Y_{...} = k + \alpha_.\beta_.$$

Thus it would follow that, in the population model,

$$\mu = k + \alpha_{\cdot}\beta_{\cdot}$$

$$a_i = (\alpha_i - \alpha_{\cdot})\beta_{\cdot}$$

$$b_j = \alpha_{\cdot}(\beta_j - \beta_{\cdot})$$

$$(ab)_{ij} = (\alpha_i - \alpha_{\cdot})(\beta_j - \beta_{\cdot}),$$

and hence the parameters of the definitional statistical model would be related by

$$a_i b_j = (\mu - k)(ab)_{ij} .$$

5. Some estimates, under general conditions

If we define

$$\hat{\mu} = x_{\cdots}$$

$$\hat{a}_i = x_{i\cdot\cdot} - x_{\cdots}$$

$$\hat{b}_j = x_{\cdot j\cdot} - x_{\cdots}$$

$$c_{ij} = (x_{ij\cdot} - x_{i\cdot\cdot} - x_{\cdot j\cdot} + x_{\cdots}) = (\hat{ab})_{ij}$$

then it is easy to check from the statistical model given in Section 1 above that these are unbiased estimates of μ, a_i, b_j, and $(ab)_{ij}$ respectively, independent of any assumption concerning functional structure.

Furthermore, if we randomly subdivide the observations into r replicates of ab observations each, then we can define

$$\hat{a}_i^f = x_{i\cdot f} - x_{\cdot\cdot f}$$

$$\hat{b}_j^f = x_{\cdot jf} - x_{\cdot\cdot f}$$

which are of course unbiased estimates of a_i and b_j respectively.

Of course the analysis of variance residual mean square

$$s^2 = \frac{1}{ab(r-1)} \sum_{ijf} (x_{ijf} - x_{ij.})^2$$

is an unbiased estimate of $(\sigma^2 + \sigma_p^2)$, as given in Table 8.

We shall use the notation V( ) to denote variance. It may be checked, using the statistical model of Section 1 above, that

$$V(\hat{\mu}) = -\sigma_p^2/P + (\sigma^2 + \sigma_p^2)/rab$$

$$V(\hat{a}_i) = (a-1)(\sigma^2 + \sigma_p^2)/rab$$

$$V(\hat{b}_j) = (b-1)(\sigma^2 + \sigma_p^2)/rab$$

$$V(c_{ij}) = (a-1)(b-1)(\sigma^2 + \sigma_p^2)/rab$$

$$V(\hat{a}_i^f) = (a-1)(\sigma^2 + \sigma_p^2)/ab$$

$$V(\hat{b}_j^f) = (b-1)(\sigma^2 + \sigma_p^2)/ab \ .$$

We now show that $\hat{a}_i$ and $\hat{b}_j$ are uncorrelated, for their covariance is

$$E(e_{i..} - e_{...})(e_{.j.} - e_{...}),$$

and

$$E(e_{i..}e_{.j.}) = E(e_{i..}e_{...}) = E(e_{.j.}e_{...}) = E(e_{...}^2)$$

$$= -\sigma_p^2/P + (\sigma^2 + \sigma_p^2)/rab \ .$$

It follows that $\hat{a}_i^f$ and $\hat{b}_j^f$ are also uncorrelated.

Hence each of

$$\hat{a}_i \hat{b}_j \text{ and } d_{ij} = \frac{1}{r} \sum_f \hat{a}_i^f \hat{b}_j^f$$

are unbiased estimates of $a_i b_j$, independent of any assumptions on

functional structure, using the statistical model of (IIA7).

Further, using only the properties of this model, we may show that the following quantities are mutually uncorrelated:

$$\hat{a}_i (\text{or } \hat{a}_i^f),\ \hat{b}_j (\text{or } \hat{b}_j^f),\ c_{ij},\ (x_{ijf} - x_{ij.}) \ .$$

6. A geometric evaluation

Combining the ideas and results of the preceding Sections 4 and 5 it is evident that a simple geometric evaluation may be made of the adequacy of the postulated multiplicative functional structure. For, if the descriptive model is indeed multiplicative, then

$$E(\hat{a}_i \hat{b}_j) = (\mu - k)\ E(c_{ij})$$

and,

$$E(d_{ij}) = (\mu - k)\ E(c_{ij}).$$

Hence a plot of the points $\left\{(\hat{a}_i \hat{b}_j,\ c_{ij})\right\}$ or $\left\{(d_{ij}, c_{ij})\right\}$, $(i = 1, 2, \ldots, a;\ j = 1, 2, \ldots, b)$ would tend to scatter around a straight line passing through the origin and having (unknown) slope $(\mu - k)$.

This rapid visual check not only gives a rough evaluation of the postulated structure but would provide a rough estimate of $(\mu - k)$ and hence of k, since $E(\hat{\mu}) = \mu$.

We note that the reasonableness of this geometric procedure is independent of any assumptions other than those built into the statistical model of Section 1.

7. Further assumptions and results

We shall now strengthen our assumptions, thereby weakening our position somewhat but simplifying the mathematics so that meaningful test criteria and distributional approximations may be suggested. The additional assumption we make is that the components

$$e_{ijf} = \sum_m \rho_m^{ijf} (p_m + \epsilon_{ijm})$$

may be treated as normally and independently distributed random variables having common mean zero and common variance

$$\sigma_o^2 = \sigma^2 + \sigma_p^2 .$$

Under these conditions, the sets of estimates

$$\{\hat{a}_i\} , \{\hat{b}_j\} , \{c_{ij}\} \text{ and } s^2$$

and $\{\hat{a}_i^f\} , \{\hat{b}_j^f\} , \{c_{ij}\}$ and s^2,

are independently distributed. Then

$$V(\hat{a}_i\hat{b}_j) = \left[a_i^2 (b-1) + b_j^2 (a-1) + \sigma_o^2 (a-1)(b-1)/rab\right] \sigma_o^2/rab .$$

$$V(d_{ij}) = \left[a_i^2 (b-1) + b_j^2(a-1) + \sigma_o^2(a-1)(b-1)/ab\right] \sigma_o^2/rab$$

$$= \frac{\sigma_o^2}{rab} v_{ij} , \text{ say.}$$

Thus $V(\hat{a}_i\hat{b}_j) < V(d_{ij})$ for all $r > 1$. However the difference in variances will usually be unimportant, for the terms $a_i^2(b-1)$ and $b_j^2(a-1)$ will dominate either variance.

The reason for considering d_{ij} at all is of course that it is the mean of r products and hence we might feel more easy* about the assumption that d_{ij} is approximately normally distributed (with mean $a_i b_j$ and variance $v_{ij}\sigma_o^2/rab$) than we would for about a corresponding assumption for $\hat{a}_i\hat{b}_j$.

Under the assumptions of this section c_{ij} is normally distributed with mean $(ab)_{ij}$ and variance $(a-1)(b-1)\sigma_o^2/rab$; and further the c_{ij} are independent of the d_{ij} (or the $\hat{a}_i\hat{b}_j$). Thus, if we can take d_{ij} as approximately normal then it would seem reasonable that

$$(d_{ij} - k_o c_{ij}), \text{ where } k_o \text{ is any constant,}$$

is approximately normally distributed with mean

$$a_i b_j - k_o c_{ij} ,$$

and variance

$$(v_{ij} + k_o^2 (a-1)(b-1))\sigma_o^2/rab .$$

Unbiased estimates of v_{ij} are

$$\hat{\hat{v}}_{ij} = \hat{a}_i^2 (b-1) + \hat{b}_j^2(a-1) + s^2 (a-1)(b-1)(r-2)/abr,$$

and

$$\hat{v}_{ij} = \frac{(b-1)}{r} \sum_f (\hat{a}_i^f)^2 + \frac{(a-1)}{r} \sum_f (\hat{b}_j^f)^2 - s^2(a-1)(b-1)/ab .$$

---

*On the other hand the use of d_{ij} is open to the objection that it may occasion different evaluation on the basis of the same set of data.

8. Distributional approximations

We have seen that under the postulated functional structure the parameters of the statistical model are related by

$$a_i b_j = (\mu - k)\, c_{ij} .$$

Hence, under this hypothesis,

$$E(d_{ij}) = (\mu - k)\, E(c_{ij})$$

even under the very broad conditions of Section 1 of this division.

Further if the assumptions and approximations of Section 7 are reasonable then, under the hypothesis of present concern,

$$(d_{ij} - (\mu - k)\, c_{ij})$$

is, approximately, normally distributed with mean 0 and variance

$$(v_{ij} + (a-1)(b-1)(\mu-k)^2)\, \sigma_o^2/rab \quad .$$

Hence, approximately,

$$\frac{(d_{ij} - (\mu - k)\, c_{ij})^2}{(v_{ij} + (a-1)(b-1)(\mu-k)^2)\sigma_o^2/rab}$$

is distributed as chi squared with one degree of freedom, under the postulated multiplicative descriptive model.

Of course the d_{ij} are mutually correlated, as are also the c_{ij}. We might approximate the joint distribution of the $\{(d_{ij}-(\mu-k)\, c_{ij})\}$ by a multivariate normal distribution, and in place of an approximate

"F" statistic, which we detail below, consider a "T^2" statistic. Such a procedure appears to be entirely unmanageable computationally for the purpose intended here and we shall not report here the results on this line which we have obtained.

If, however, we can treat the $(d_{ij} - (\mu-k)\, c_{ij})$ as independent then, approximately,

$$\frac{rab}{\sigma_o^2} \sum_{ij} (d_{ij} - (\mu-k)\, c_{ij})^2/(v_{ij} + (a-1)(b-1)(\mu-k)^2)$$

is distributed as chi squared with ab degrees of freedom.

Hence, approximately,

$$F_o(\mu-k) = \frac{r}{s^2} \sum_{ij} (d_{ij} - (\mu-k)\, c_{ij})^2/(v_{ij} + (a-1)(b-1)(\mu-k)^2)$$

has an F distribution with ab and ab(r-1) degrees of freedom.

The function F_o is of course not of much use to us directly since v_{ij} involves the unknown parameters a_i, b_j and σ_o^2. We have however an unbiased estimate of v_{ij}, namely $\hat{v}_{ij}$ (or $\hat{\hat{v}}_{ij}$) given in Section 7. Thus we put

$$\hat{F}_o(\mu-k) = \frac{r}{s^2} \sum_{ij} (d_{ij} - (\mu - k)\, c_{ij})^2/(\hat{v}_{ij} + (a-1)(b-1)(\mu-k)^2) .$$

An adjustment of degrees of freedom may give a more reasonable approximation.

9. Uses of $\hat{F}_o$

From its structure the function $\hat{F}_o$ would appear to be sensitive to departures from the postulated functional structure in the sense that

then no values which $(\mu-k)$ might take on would make $\hat{F}_o$ "small". On the other hand, if the postulated structure is reasonable, then for some values of $(\mu-k)$ $\hat{F}_o(\mu-k)$ will be "small".

Given a hypothesized value for k, say k°, in the postulated descriptive model

$$Y_{ij.} = k + \alpha_i\beta_j,$$

we could calculate $\hat{F}_o(\hat{\mu}-k^\circ)$ and compare with the F distribution for a significance test, large values of $\hat{F}_o$ casting doubt on the hypothesis under test. In particular we might test $k = 0$.

A much less restrictive and more useful application would be to determine the range of values (if any do in fact exist) of $(\mu-k)$ for which $\hat{F}_o(\mu-k)$ takes on values less than, say, the 5% point of the F distribution. If there were no doubt as to such a range being continuous, then such a calculated interval would be an approximate 95% confidence interval for $(\mu-k)$; and the existence of such values of $(\mu-k)$ would be statistical evidence that the postulated functional structure might be adequate for the situation. An interval estimate for k might then be obtained using $\hat{\mu}$.

To find an interval estimate for $(\mu-k)$ we proceed by finding a "reasonable" preliminary estimate for $(\mu-k)$ and then plot $\hat{F}_o(\mu-k)$ in the neighborhood of this value to give upper and lower limits. This bypasses the possible problem of a discontinuous range.

A "reasonable" starting value for $(\mu-k)$ can be obtained in a number of ways. In some cases $\hat{\mu}$ might be a reasonable place to start. Other-

wise we would use the linear regression of d_{ij} on c_{ij}, either by eye estimate of slope of a scatter diagram or by using

$$\sum_{ij} d_{ij} c_{ij} / \sum_{ij} c_{ij}^2 .$$

The regression of d_{ij} on c_{ij} is to be preferred because the latter will, in general, have the smaller variance.

IV. SUMMARY

Our concerns in this thesis have centered about the derivation and interpretation of linear statistical models for randomized experiments and with their application and use in the analysis of such experiments. The central features of both the underlying philosophy and methodological detail were the concepts of "true response", "experimental unit" and the use of randomization in the design and procedure.

We have given in Part I some introductory or expository discussions on experimental error, randomization, classification of types of models, the analysis of variance and the relation of some aspects of this thesis to other published and unpublished work.

In Part II we have studied the standard statistical experimental designs, namely the completely randomized, randomized block, latin square and split plot designs, under quite general experimental conditions. Linear statistical models, whose components depend on parameters defined on the population of "true" responses and on certain "dummy" random variables which reflect the (randomized) experimental design and procedure, have been derived for each of these designs under the general experimental conditions. These models, which are definitional in the sense that they do not depend on notions of causality, mechanism or functional relation, have been employed in the study of the analysis of variance and of various estimation questions. In

particular, general results on expectations of analysis of variance mean squares are given, which are valid for so-called fixed, mixed and random model situations as special cases as well as for all intermediates. Interactions of treatments with experimental units were found to occasion a "bias" in the analysis of variance in the sense that one cannot, in general, obtain from the mean squares unbiased estimates of components of variation. The estimable (under all conditions) functions in the analysis of variance were found to have well-determined and defined structures, and considerable attention was given to describing and detailing these structures. The pattern of expected mean squares in terms of these estimable functions was found to be quite simple and this pattern was of very considerable value in unification, simplification and extension of the results. Considerable discussion was given of the general implicit and explicit definitions of the estimable functions in terms of the components of variation and population sizes, and of the pattern of expected mean squares for general situations in terms of these estimable functions. Either as part of overall discussions or in connection with the study of particular situations attention has been given to: general expected mean squares for cases of unequal and proportional numbers in the cells; alternative analyses of variance for certain unbalanced experiments; the dependence of reasonable assumptions concerning properties of and relationships among parameters of the conceptual population on the intent of the experimental design; questions concerning estimates of effects, of errors, and of components of variation; the selection of appropriate

test criteria for various null hypotheses; the general effects of non-additivities; the accomplishments of randomization; the dependence of "error terms" on the experimental procedure. The basic assumptions underlying the development have been summarized, and their weakening discussed, in a section under general discussions.

In Part III we have discussed briefly the notion of functional structure and the problem of elucidation of functional structure by statistical techniques. The (meager) literature of relevance on this matter was reviewed in the context. The idea that the underlying functional structure determines certain properties of and relationships among components of the definitional (derived) linear statistical model was put forth. This idea was applied in the derivation of an approximate test for a multiplicative functional structure. A rough visual method for the evaluation of the adequacy of the postulated functional structure was described. A more objective basis for evaluation and for the interval estimation of an unknown parameter was developed, based on a number of statistical assumptions and distributional (mathematical) approximations.

V. REFERENCES

Anderson, R. L. and T. A. Bancroft (1952). Statistical theory in research. McGraw-Hill Book Co., New York.

Anscombe, F. L. (1948). The validity of comparative experiments. J. Roy. Stat. Soc. Series A. 111:181-200.

_____ and J. W. Tukey (1954). Transforming counted observations. Paper presented at meeting of Amer. Stat. Assoc. and Biometrics Soc., Montreal.

Balmukand, B. H. (1928). Studies in crop variation. V. The relation between yield and soil nutrients. J. Agr. Sci. 18:602-625.

Bancroft, T. A. (1944). On biases in estimation due to the use of preliminary tests of significance. Ann. Math. Stat. 15:190-204.

Barnard, G. A. (1952). Discussion on a paper by Tocher. J. Roy. Stat. Soc. 14:91-92.

Bennett, C. A. and N. L. Franklin (1954). Statistical analysis in chemistry and the chemical industry. John Wiley and Sons, Inc., New York.

Berkson, J. (1950). Are there two regressions? J. Amer. Stat. Soc. 45:164-180.

Bozivich, Helen, T. A. Bancroft and H. O. Hartley (1955). Preliminary tests of significance and pooling procedures. Part one. Analysis of variance procedures for pooling mean squares for certain incompletely specified models. Unpublished final report to Wright Air Development Center. Iowa State College Statistical Laboratory.

Cochran, W. G. (1947). Some consequences when the assumptions for the analysis of variance are not satisfied. Biometrics. 3:22-38.

_____ (1951). Testing a linear relation among variances. Biometrics. 7:17-32.

_____ and G. M. Cox (1950). Experimental designs. John Wiley and Sons, Inc., New York.

Cornfield, J. (1944). On samples from finite populations. Amer. Stat. Assoc. 39:236-239.

_____(1953). The two-way classification in the analysis of variance from the point of view of urn sampling. Office of Biometry, National Institute of Health. (Mimeo. rept.)

Cramér, H. (1946). Mathematical methods of statistics. Princeton University Press. Princeton.

Crump, S. L. (1951). The present status of variance component analysis. Biometrics. 1:1-17.

Davies, O. L. (editor) (1954). The design and analysis of industrial experiments. Oliver and Boyd, London.

Eden, T. and F. Yates (1933). On the validity of Fisher's z-test when applied to an actual example of non-normal data. J. Agr. Sci. 23:6-16.

Eisenhart, C. (1947). The assumptions underlying the analysis of variance. Biometrics. 3:1-21.

Fisher, R. A. (1918). The correlation between relatives on the supposition of Mendelian inheritance. Trans. Roy. Soc. Edin. 52:339-433.

_____(1924). On a distribution yielding the error functions of several well known statistics. Proc. Int. Math. Cong., Toronto.

_____(1926). The arrangement of field experiments. J. Min. Agr. Eng. 33:503-513.

_____(1935a). The design of experiments. Oliver and Boyd, London.

_____(1935b). Discussion on a paper by J. Neyman et al. J. Roy. Stat. Soc. Suppl. 2:155-157.

_____(1936). The coefficient of racial likeness and the future of craniometry. J. Roy. Anthrop. Inst. of Gr. Brit. and N. Ireland. 66:57-63.

_____(1950). Statistical methods for research workers, 11th edition. Hafner Publishing Co., New York.

_____(1954). The analysis of variance with various binomial transformations. Biometrics. 10:130-139.

_____ and W. A. Mackenzie (1923). Studies in crop variation. II. The manurial response of different potato varieties. J. Agr. Sci. 13:311-320.

_____ and F. Yates (1938). Statistical tables. Oliver and Boyd, London.

Graybill, F. (1954). Variance heterogeneity in a randomized block design. Biometrics. 10:516-520.

Greenberg, B. G. (1951). Why randomize? Biometrics. 7:309-322.

Grundy, P. M. and M. J. R. Healy (1950). Restricted randomization and quasi-latin squares. J. Roy Stat. Soc. B. 12:286-291.

Hald, A. (1952). Statistical theory with engineering applications. John Wiley and Sons, Inc., New York.

Hoel, P. G. (1954). Introduction to mathematical statistics. Second edition. John Wiley and Sons, Inc., New York.

Hooke, R. (1953). Sampling from a matrix, with applications to the theory of testing. Memorandum Report 53. Princeton University. (Mimeo. rept.)

_____ (1954a). Moments of moments in matrix sampling - an extension of polykays. Memorandum Report 55. Princeton University. (Mimeo. rept.)

_____ (1954b). The estimation of polykays in the analysis of variance. Memorandum Report 56. Princeton University. (Mimeo. rept.)

Hotelling, H. (1947). Multivariate q uality control. Selected techniques of statistical analysis. Chapt. 3. McGraw-Hill Book Co., New York.

Huntsberger, D. (1955). Preliminary tests of significance. Part two. A generalization of a preliminary testing procedure. Unpublished final report to Wright Air Development Center. Iowa State College Statistical Laboratory.

Irwin, J. O. (1931). Mathematical theorems involved in the analysis of variance. J. Roy. Stat. Soc. 94:284-293.

Kempthorne, O. (1952a). The design and analysis of experiments. John Wiley and Sons Inc., New York.

_____ (1952b). The randomization theory of experimental inference. Paper given before meeting of Amer. Stat. Assoc. and Biometrics Soc., Chicago. In press, J. Amer. Stat. Assoc. (To appear in 1955).

_____ and W. D. Barclay (1953). The partition of error in randomized blocks. J. Amer. Stat. Assoc. 48:610-614.

Kendall, M. G. (1946). The advanced theory of statistics, volume II. First edition. Hafner Publishing Co., New York.

_____ (1951). Regression structure and functional relationship. Biometrika. 38:11-25.

McCarthy, M. D. (1937). On the application of the z-test to randomized blocks. Ann. Math. Stat. 10:337-359.

Mentzer, E. G. (1953). Tests by the analysis of variance. Wright Air Development Center Technical Report 53-23.

Mood, A. M. (1950). Introduction to the theory of statistics. McGraw-Hill Book Co., New York.

Moore, P. G. and J. W. Tukey (1954). Answer to Query. Biometrics. 10:562-568.

Neyman, J., with K. Iwaszkiewiez and St. Kolodziejczyk (1935). Statistical problems in agricultural experimentation. J. Roy. Stat. Soc. Suppl. 2:107-154.

Pearce, S. C. (1953). Field experimentation with fruit trees and other perennial plants. Technical communication No. 23 of the Commonwealth Bureau of Horticulture and Plantation Crops, East Malling, Maidstone, Kent.

Pitman, E. J. G. (1937). Significance tests which may be applied to samples from any population. III. The analysis of variance test. Biometrika. 29:322-335.

Quenouille, M. H. (1953). The design and analysis of experiment. Charles Griffin and Co., Ltd., London.

Rafferty, J. A. (1950). Mathematical models in biological theory. Amer. Scientist. 38:549-567.

Rao, C. R. (1952). Advanced statistical methods in biometric research. John Wiley and Sons, Inc., New York.

Robbins, H. and E. J. G. Pitman (1949). Application of the method of mixtures to quadratic forms in normal variates. Ann. Math. Stat. 20:552-560.

Satterthwaite, F. E. (1946). An approximate distribution of estimates of variance components. Biometrics. 2:110-114.

Scheffé, H. (1954). Statistical methods for evaluation of several sets of constants and several sources of variability. Chem. Eng. Progress. 50:200-205.

_____ (1955). A "mixed model" for the analysis of variance. Univ. of California. (Mimeo. rept.)

Tick, L. (1954). New York, N.Y. (Private communication to O. Kempthorne, Ames, Iowa).

Tolman, R. C. (1938). The principles of statistical mechanics. Oxford at the Clarendon Press.

Tukey, J. W. (1949a). One degree of freedom for non-additivity. Biometrics. 5:232-242.

_____ (1949b). Dyadic anova, an analysis of variance for vectors. Human Biology. 21:65-110.

_____ (1949c). Interaction in a row by column design. Memorandum Report 18, Princeton University. (Mimeo. rept.)

_____ (1954). Causation, regression and path analysis. Statistics and mathematics in biology, Chapt. 3. Iowa State College Press, Ames.

_____ (1955). Answer to Query. Biometrics. 11:111-113.

Ward, G. C. and I. D. Dick (1952). Non-additivity in randomized block designs and balanced incomplete block designs. N. Z. J. Sci. and Tech. 33:430-435.

Welch, B. L. (1937). On the z-test in randomized blocks and latin squares. Biometrika. 29:21-52.

Wilk, M. B. (1953a). The randomization analysis of a generalized randomized block design. Unpublished M. S. Thesis. Iowa State College.

_____ (1953b). The logical derivation of linear models for experimental situations and their use in selecting the appropriate error term in the analysis of variance. I. Analysis of Variance Project Report AV-3, Iowa State College Statistical Laboratory. (Mimeo. rept.)

_____ (1954a). The logical derivation of linear models... II. On the analysis of a general randomized block type design. Analysis of Variance Project Report AV-4, Iowa State College Statistical Laboratory. (Mimeo. rept.)

_____ (1954b). The logical derivation of linear models... V. The expectations of mean squares for a general three-way classification (part 2). Analysis of Variance Project Report AV-9. Iowa State College Statistical Laboratory. (Mimeo. rept.)

_____ (1954c). The logical derivation of linear models... VI. Interactions with experimental units. Analysis of Variance Project Report AV-10. Iowa State College Statistical Laboratory. (Mimeo. rept.) Presented before a meeting of Inst. Math. Stat., Iowa City.

_____ (1955). The randomization analysis of a generalized randomized block design. Biometrika. 42:1-10.

_____ and O. Kempthorne (1954a). The logical derivation of linear models... III. The expectation of mean squares for a general three-way classification. Analysis of Variance Project Report AV-7. Iowa State College Statistical Laboratory. (Mimeo. rept.)

_____ and _____ (1954b). The logical derivation of linear models... IV. Fixed, mixed and random models. Analysis of Variance Project Report AV-8. Iowa State College Statistical Laboratory. (Mimeo. rept.) Presented before meeting of Amer. Stat. Assoc., Montreal.

_____ and _____ (1955). Derived linear models and their use in the analysis of randomized experiments. Unpublished final report to Wright Air Development Center. Iowa State College Statistical Laboratory.

Yates, F. (1932). The principles of orthogonality and confounding in replicated experiments. J. Agr. Sci. 23:108-145.

_____ (1936). Incomplete latin squares. J. Agr. Sci. 26:301-315.

_____ (1937). The design and analysis of factorial experiments. Imp. Bur. Soil Sci. Harpenden, England.

_____ (1939). The comparative advantages of sy stematic and randomized arrangements in the design of agricultural and biological experiments. Biometrika. 30:440-466.

_____ and W. G. Cochran (1938). The analysis of groups of experiments. J. Agr. Sci. 28:556-580.

VI. ACKNOWLEDGEMENTS

It is a pleasure to acknowledge the assistance of Professor O. Kempthorne in the preparation of this thesis and in the investigations underlying it.

Part of the investigations on which the thesis is based were supported by Wright Air Development Center under Contract No. AF33(616)-2145.

Heartfelt thanks are due to Mrs. Margaret Kirwin for her invaluable assistance and advice in the preparation of the manuscript.